Libros de Cátedra

Introducción a la Química Medicinal

Luciana Gavernet (coordinadora)

FACULTAD DE
CIENCIAS EXACTAS

e
exactas





INTRODUCCIÓN A LA QUÍMICA MEDICINAL

Luciana Gavernet

(coordinadora)

Facultad de Ciencias Exactas







Al Prof. Dr. Luis Bruno Blanch.

Agradecimientos

Los autores de este libro queremos agradecer al Profesor Extraordinario Dr. Luis Bruno-Blanch por su contribución a la formación en Química Medicinal de los alumnos de la carrera de Farmacia de la Facultad de Ciencias Exactas (UNLP) durante ininterrumpidos 40 años dedicados a la docencia universitaria. Los continuos esfuerzos del Dr. Bruno-Blanch constituyen un aporte fundamental para el desarrollo de las Ciencias Químicas y Farmacéuticas tanto a nivel científico, como a nivel de formación de docentes-investigadores. Su devoción a la formación de recursos humanos se ha traducido, recientemente, en la creación del Laboratorio de Investigación y Desarrollo de Bioactivos (LIDeB) de la UNLP.

Índice

Prólogo

Agradezco la oportunidad que me brindan la responsable y los diferentes autores de este, el primer libro de Química Medicinal de la Carrera de Farmacia de la Facultad de Ciencias Exactas de la UNLP, en dirigirles unas palabras además de permitirme expresar el significado que tiene para mí la concreción del mismo.

Su utilidad, por los diferentes contenidos y la profundidad de cada tema, no solo será materia de lectura para los alumnos de la Carrera de Farmacia sino, será de lectura de todos aquellos que estén interesados en conocer las diferentes metodologías que hace uso hoy la Química Medicinal, describiendo las áreas que contribuyen para la obtención de nuevos fármacos.

Nuestra Facultad incorpora esta área temática por primera vez en su programa de 1979, con el nombre de Farmacoquímica, y luego con el nombre de Química Medicinal, y del cual formé parte del cuerpo docente desde su inicio. Siempre se planteó que era necesario de que nuestros alumnos tuvieran un material acorde a los conocimientos actuales y que fuera una puerta que les permitiera ver una perspectiva del Farmacéutico en el Diseño de Nuevos Fármacos, construyendo nuevas moléculas para contribuir a mejorar las diferentes patologías que afectan a nuestra sociedad.

Desde el inicio de su lectura, uno puede apreciar que es un libro escrito por gente formada en distintas disciplinas, en general por jóvenes y excelentes investigadores integrados en un proyecto común, que se mueven con soltura, actualidad y profundidad en los temas que cada uno desarrolla.

Todos abordan, como objetivo fundamental, el descubrimiento de nuevos fármacos con un profundo conocimiento de las metodologías que explican y usan, tratando de trasmitir, con ejemplos y conclusiones, que los lectores pueden sentirse participes activos de la creación de nuevos conocimientos.

Esa creación, las de nuevas moléculas biológicamente activas, es un arte que necesita vocación, espíritu joven y creativo, movido por un gran deseo de que se haga realidad su obra para llegar como medicamento a un paciente que lo necesita para mejorar su calidad de vida.

Deseo que este libro, en su primera edición y en las que seguro vendrán, permita al lector estudiante de Farmacia, comprender y participar de manera activa en el curso, y ese conocimiento pueda transformarlo en formación profesional que le permita acceder a la atención farmacéutica en sus distintas expresiones, como a la investigación científica en el desarrollo de nuevos medicamentos.

Prof. Dr. Luis E. Bruno Blanch

La Plata, noviembre de 2018

INTRODUCCIÓN
Química Medicinal, la importancia en la formación del profesional farmacéutico

Carolina L. Bellera, Mauricio E. Di Ianni y Luciana Gavernet

La **Química Medicinal** o **Química Farmacéutica** tiene como objetivo el estudio de los fármacos desde el punto de vista químico, así como de los principios básicos de su diseño. Rigurosamente, la IUPAC interpreta que la Química Medicinal comprende el descubrimiento, el desarrollo, la identificación e interpretación del modo de acción de los compuestos biológicamente activos a nivel molecular. Se pone especial énfasis en las drogas, pero los intereses de los químicos medicinales no solo se restringen a estas estructuras, sino que incluye compuestos bioactivos en general. Además, considera el estudio, la identificación y la síntesis de productos metabólicos de las drogas y compuestos relacionados. Esta disciplina posee profundas raíces en la química, con alcance hacia la biología, medicina y ciencias farmacéuticas. Por lo tanto, las áreas a abordar tienen características netamente interdisciplinarias, como la química orgánica, la fitoquímica, la farmacología, la toxicología, la biología molecular, la bioquímica, la química física, la informática y las matemáticas. Asimismo, incluye el estudio de drogas existentes, sus propiedades farmacológicas, la interpretación de sus mecanismos de acción, efectos tóxicos y las relaciones existentes entre la estructura química y la actividad biológica. Por lo tanto, la enseñanza de la Química Medicinal como asignatura de grado de la carrera de Farmacia resulta fundamental para el desarrollo del farmacéutico, a fin de proporcionarle las herramientas necesarias para su capacitación y actualización en las metodologías asociadas al descubrimiento de un nuevo compuesto bioactivo.

Algunas definiciones importantes

Como consecuencia del elevado grado de interacción entre disciplinas que posee la Química Medicinal, es necesario definir un vocabulario específico para facilitar el estudio y la comprensión del contenido del libro. Para ello, el estudiante debe tener presente la terminología detallada a continuación.

Ligando: Sustancia capaz de desencadenar, potenciar, inhibir o modular una respuesta biológica tras unirse específica o inespecíficamente a una biomolécula (proteínas, lípidos o ácidos nucleicos).

Blanco Molecular: Biomolécula susceptible de interaccionar con un ligando y experimentar cambios que encaucen una respuesta biológica determinada.

Receptor: Proteína susceptible de interaccionar con un ligando y experimentar cambios que encaucen una respuesta biológica determinada.

Droga: Toda sustancia exógena biológicamente activa.

Prodroga o profármaco: Sustancia farmacológica que se administra en forma inactiva y que ejerce su efecto luego de ser metabolizado por el organismo.

Principio Activo: Sustancia química o mezcla de sustancias relacionadas, de origen natural o sintético que, poseyendo un efecto farmacológico específico, se emplea en medicina humana. También nos referimos a él como *fármaco*, *ingrediente activo* o *ingrediente farmacológicamente activo (IFA)*. Habitualmente el principio activo es aludido como "droga" aunque en las ciencias farmacéuticas preferimos la terminología previa.

Medicamento: Cuando se combinan el o los principios activos con ingredientes inactivos o excipientes siguiendo determinados procesos de manufactura dan lugar al medicamento. El conjunto de excipientes constituye el vehículo del principio activo, y no poseen actividad farmacológica intrínseca. Si bien los componentes del vehículo no presentan actividad farmacológica per se, su composición influye en el resultado del tratamiento farmacológico. Entre otras cosas, los excipientes son determinantes en los procesos de liberación, absorción, distribución, metabolismo y excreción (LADME) y por ende de la biodisponibilidad (conceptos estudiados en Biofarmacia), la aceptabilidad del tratamiento por parte del paciente y la estabilidad del principio activo en el vehículo. Diferentes tipos de excipientes se estudian en disciplinas como Farmacotecnia o Tecnología Farmacéutica.

Especialidad Medicinal: Todo medicamento, designado por un nombre convencional, sea o no una marca comercial de un laboratorio farmacéutico, o por el nombre genérico que corresponda a su composición y contenido, preparado y envasado uniformemente para su distribución y expendio, de composición cuantitativa definida declarada y verificable, de forma farmacéutica estable y acción terapéutica comprobable.

Droga "Me Too": Fármacos estructuralmente muy similares a otros ya conocidos, únicamente diferenciados por variaciones menores en el perfil farmacológico.

Compuestos Hits: Moléculas que demuestren actividad satisfactoria frente al blanco molecular. Posteriormente, los hits son optimizados para mejorar su potencia y otras propiedades significativas a través de ciclos de optimización, para conducir a los denominados "Compuestos líderes".

Compuesto Líder (*Lead compound*): Compuesto químico que manifiesta una actividad determinada NO superada por otros compuestos y cuyos cambios estructurales realizados hasta el momento no han mejorado su actividad ni disminuido su toxicidad.

Modelado Molecular: Término que engloba a todo método teórico y/o computacional usado para describir o imitar el comportamiento de moléculas.

Docking: Metodología computacional de modelado molecular para la exploración de los posibles modos de unión de un ligando a un determinado blanco molecular.

Actividad biológica: Todo efecto benéfico, terapéutico, adverso o tóxico de una determinada droga sobre la materia viva.

Cribado Virtual (Tamizado Virtual o Screening Virtual): Conjunto heterogéneo de metodologías computacionales que permiten explorar la representación digital de las entidades químicas almacenadas en grandes repositorios de estructuras con el propósito de seleccionar aquellas que se predicen con la actividad biológica buscada.

Luego de esta breve introducción, los invitamos a transitar las siguientes páginas del libro, donde se describirán las estrategias generales vinculadas al origen y descubrimiento de fármacos (Capítulos 1 y 2) y se detallarán metodologías para la síntesis química de dichos compuestos (Capítulo 3). Luego se explicarán las herramientas computacionales útiles para la búsqueda o el diseño racional de compuestos activos (Capítulos 4 a 6) y se enumerarán algunas estrategias de optimización de estructuras para mejorar la llegada de los fármacos al sitio de acción (Capítulo 7). Finalmente, se analizarán diferentes metodologías de evaluaciones biológicas a nivel preclínico utilizadas para el descubrimiento de fármacos (Capitulo 8).

Es nuestro deseo como docentes de la Cátedra de Química Medicinal que el contenido del libro de cátedra resulte no solo interesante para la formación general de futuros farmacéuticos sino también para clarificar dudas que puedan surgir en el período de aprendizaje de los estudiantes de la carrera de farmacia de la UNLP.

CAPÍTULO 1
Descubrimiento y desarrollo de fármacos

Carolina L. Bellera y Mauricio E. Di Ianni

1.1. Descubrimiento de fármacos

El proceso de diseño y desarrollo de fármacos postulado por la química medicinal posee tres etapas fundamentales y consecutivas:

Descubrimiento: es el paso inicial o fase cero del proceso. Comprende la identificación de nuevas sustancias, denominadas *hits, prototipos activos o cabezas de serie*, que manifiesten una actividad biológica determinada. Estos compuestos pueden ser obtenidos, entre otros métodos, por cribado sistemático de una gran cantidad de sustancias de origen sintético o de fuentes naturales como plantas, animales o microorganismos.

Optimización: esta etapa busca mejorar la estructura cabeza de serie. El proceso de optimización apunta principalmente al aumento de la potencia y selectividad, y a la disminución de la toxicidad del compuesto. Es característico de este paso el establecimiento y análisis de las relaciones entre la estructura química y la actividad biológica (SARs, del inglés *Structure Activity Relationships*) las cuales son la base de las modificaciones estructurales racionales que conducirán a la optimización de la estructura. La prueba y error es la otra cara de la moneda, filosóficamente opuesta a la anterior pero históricamente utilizada como método para alcanzar la optimización de un prototipo activo.

Desarrollo: tras lograr el éxito en el paso anterior, la nueva entidad molecular (NEM) entra en un largo y costoso periodo que abarca desde el escalado de la producción de la sustancia y por ende la puesta a punto de su método de obtención, hasta las pruebas y validaciones biológicas preclínicas y clínicas que terminarán de otorgarle a la droga el carácter de medicamento con su correspondiente garantía de seguridad y efectividad para el uso medicinal propuesto.

La figura 1.1 muestra un esquema general del proceso clásico de descubrimiento de fármacos. A pesar de que este modelo se sigue aplicando con éxito, no necesariamente es el más eficiente. Podemos observar que introducir un nuevo fármaco en el mercado es un proceso que requiere varios años y es económicamente muy costoso. Como se detallará más adelante en apartado 1.3, el costo final total es difícil de estimar porque dependerá del tipo de droga, proceso etc., pero en todos los casos es muy elevado. Informes de los últimos años han reportado que introducir un fármaco novedoso puede costar hasta 2,5 billones de dólares. El tiempo promedio que se requiere desde el descubrimiento del principio activo hasta el lanzamiento

oscila entre los 12 y 15 años. Adicionalmente, existe un riesgo muy grande de fracaso. Por ejemplo, las bibliotecas de compuestos de screening iniciales (entre 10^4-10^6 compuestos), conducen a una única estructura que tiene alrededor de solo 8% de probabilidad de superar exitosamente las pruebas clínicas.

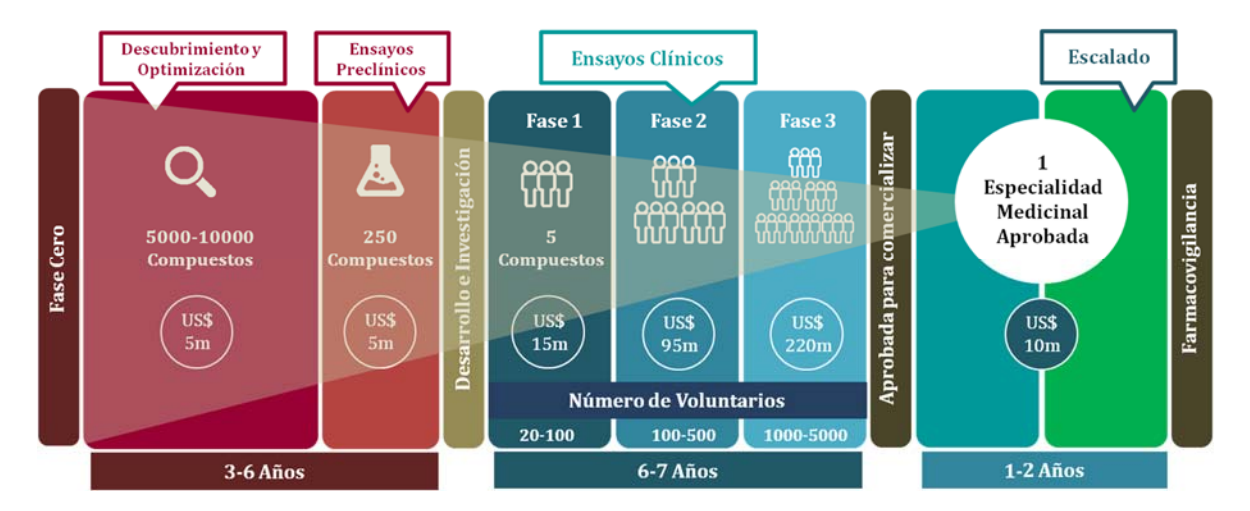


Figura 1.1: Esquema general del proceso de descubrimiento y desarrollo de un fármaco. Adaptado de Trac Services Regulatory Blog. The Drug Development Process (www.tracservices.co.uk).

1.2. Un objetivo, dos Enfoques:

El rol de la Química Medicinal en la etapa de descubrimiento de fármacos es trascendental. En esta etapa podemos diferenciar dos estrategias generales de descubrimiento de fármacos (Figura 1.2):

El enfoque conocido como ***cribado fenotípico*** o ***primero la función*** (en inglés, *function first*; equivalente en genómica con el mecanismo “hacia adelante o directo”). Es la estrategia clásica en el descubrimiento de fármacos, según la cual se comienza con un organismo que muestra un determinado fenotipo asociado a una enfermedad y se ensayan distintos fármacos candidatos para identificar aquellos eficaces y no tóxicos. En esta aproximación no es necesario postular (o dilucidar) un mecanismo de acción para desarrollar un fármaco exitoso; eventualmente este paso de identificación del mecanismo de acción podría devenir (mucho) más adelante.

El enfoque denominado ***cribado basado en el blanco terapéutico*** o ***primero el blanco*** (en inglés, *target first*; equivalente en genómica con el mecanismo “de reversa o hacia atrás”). En parte como resultado de los proyectos de secuenciación del genoma y de dilucidación del proteoma —en los cuales se identifican genes y productos génicos asociados a distintas enfermedades— surge un nuevo enfoque que se basa en la capacidad de los fármacos de modular blancos terapéuticos específicos, a bajas concentraciones, en ensayos *in vitro* que a posteriori son validados mediante ensayos *in vivo*. A diferencia del cribado fenotípico, aquí se parte de un cri-

bado destinado a seleccionar aquellos compuestos que se unen selectivamente al blanco molecular de interés; en el caso del cribado virtual, se podrá utilizar un modelo de la estructura 3D de la proteína blanco o ligandos de referencia asociados a un mecanismo de acción específico para explorar las bases de datos. Ambas estrategias presentan ventajas y desventajas. Las fortalezas del enfoque “primero el blanco terapéutico” incluyen la capacidad de aplicar el conocimiento químico y molecular para investigar las hipótesis moleculares específicas, y la mayor viabilidad de aplicar el cribado a grandes bases de datos de moléculas pequeñas. Adicionalmente, los avances en la quimioinformática y bioinformática contribuyen enormemente al diseño, selección u optimización de los candidatos a fármacos para su interacción con un blanco terapéutico. Una desventaja de esta aproximación es que no se conoce a priori la capacidad de los fármacos así identificados para modificar la progresión de una enfermedad o controlar su sintomatología, pudiendo observarse frecuentemente una ausencia de correlación entre la actividad *in vitro* de los fármacos candidatos y su actividad *in vivo*.

En el “cribado fenotípico”, en cambio, la actividad del candidato puede extrapolarse con mayor eficacia al impacto terapéutico en un estado determinado de enfermedad. La debilidad del enfoque es la falta de una relación clara entre el mecanismo de acción y el efecto biológico, siendo común que aquél se desconozca total o parcialmente. Un desafío explorado recientemente es el de incorporar de manera efectiva nuevas tecnologías de cribado enfocados en el fenotipo.

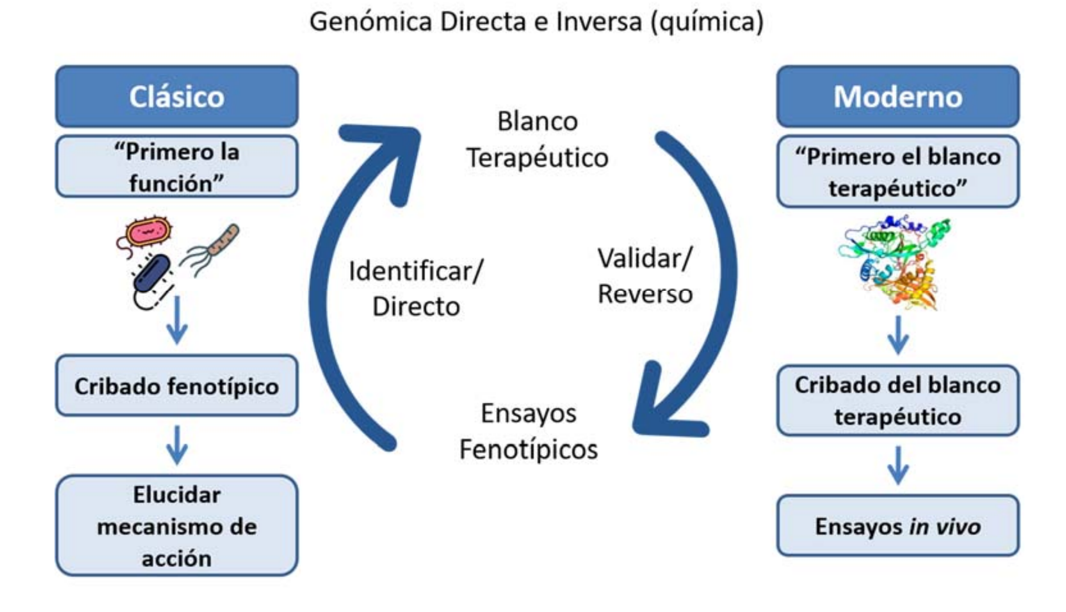


Figura 1.2: Esquema de estrategias aplicadas en el descubrimiento de fármacos.

1.3. Costos y fracasos en el desarrollo de fármacos

Las estimaciones sobre el costo medio vinculado al desarrollo de fármacos según reportes de publicaciones analizadas por Morgan y colaboradores varían desde 92 millones de dólares de costo directo (161 millones de dólares de costo de capital) hasta 883,6 millones (1,8 billones de

costo de capital). Las diferencias en las bases de datos examinadas para realizar estas estimaciones, las metodologías y la diversidad de categorías terapéuticas analizadas conducen a la amplia variación en el costo estimado. Las cifras más elevadas corresponden al análisis de información confidencial que proveen los laboratorios farmacéuticos, quienes presentan intereses respecto a la publicación de altas cifras para justificar el valor de sus productos en el mercado y para argumentar a favor de la extensión de las patentes de productos farmacéuticos.

A pesar de las diferencias en la estimación de los costos, estos números reflejan la importante inversión que insumen los estudios necesarios para que los candidatos se puedan transformar en medicamentos seguros y eficaces disponibles para su uso clínico.

El proceso de desarrollo está caracterizado por una alta tasa de fracaso. Las estimaciones de la tasa de éxito de los medicamentos que entran en las fases clínicas varían desde 11,7% a 24,0%; es decir, sólo uno o dos de cada diez proyectos que ingresan a estudios clínicos logran acceder al mercado farmacéutico bajo la forma de un medicamento innovador.

Alrededor de un 27% de los fracasos de los candidatos a fármacos en las fases preclínica y clínica se deben a falta de eficacia, y más de un 30% a problemas relacionados con la seguridad/toxicidad. En la Figura 1.3 se presenta la distribución de frecuencias de las causas de fracasos de aquellos candidatos que llegan a la etapa de evaluación clínica.

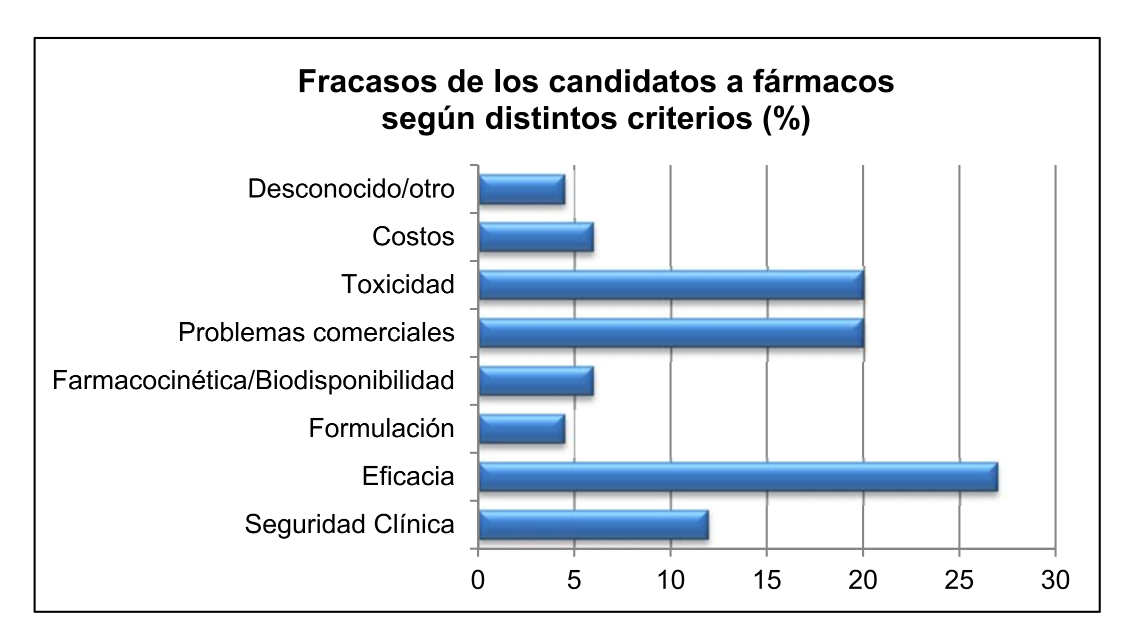


Figura 1.3: Distribución de frecuencias de causas de fracaso de los candidatos a fármacos en fase clínica.

Un análisis del año 2014 realizado por el laboratorio farmacéutico AstraZeneca, ha demostrado que en el periodo 2001-2010, de entre 198 candidatos a fármacos de dicho laboratorio, 48 (24%) fallaron por razones de seguridad entre el inicio de los estudios de toxicidad y la entrada a la fase I. De los 48 fracasos el 50% estaba relacionado a problemas de toxicidad en órganos (detectados por histopatología) o por muertes inexplicables durante las 2 semanas posteriores al inicio de la exposición.

En la Tabla 1.1 se detallan los problemas de toxicidad detectados para los 48 candidatos que no lograron avanzar en las fases clínicas. Los datos anteriores revelan que, si bien la falta de eficacia es una causa importante de atrición, se encuentra prácticamente a la par de la falta de seguridad de los candidatos estudiados.

Es por ello que uno de los objetivos relevantes para los laboratorios farmacéuticos, y para todos los profesionales químicos o farmacéuticos que trabajan en las etapas tempranas del descubrimiento de fármacos es mejorar la capacidad para predecir la probabilidad de fracaso asociada a la toxicidad antes de ingresar a las fases clínicas, dado que tomar la decisión de avanzar hacia la fase clínica representa una significativa inversión financiera y de recursos de investigación y desarrollo (I + D).

Tabla 1.1: Fracasos de fármacos debido a diferentes toxicidades.

Tipos de Toxicidad de candidatos a fármacos que han fracasado		
Predicción	**Número de Candidatos a Fármacos**	**Toxicidad**
Detectado	2 1 1 1 4 2 2 3 1 1 1 1 4 24 (50%)[a]	Hígado Gastrointestinal Sistema Nervioso Central Riñón Patologías cardiovasculares Pulmón Muerte por causa inespecífica Músculo Páncreas Tiroides Lenticular Toxicidad múltiple en órganos específicos Función cardiovascular
No detectado luego de 2 semanas	1 1 1 1 4 (8%)	Patología Cardiovascular Sistema Nervioso Central Pulmonar Reprotoxicidad
Retirado por razones que no involucran la toxicidad	9 (19%)	No se debe a toxicidad
Incierto	3 1 2 1 1 1 2 11 (23%)	Riñón Hígado Pulmón Músculo Testículo Glándula Suprarrenal Toxicidad en múltiples órganos
Total	48 (100%)	

Nota: [a] Este 50% comprende a los 20 compuestos (40%) retirados por causas atribuibles a toxicidades en órganos más los 4 (10%) compuestos debido a los riesgos cardiovasculares agudos. Adaptado de Roberts et al, 2014.

1.4. Cambio de paradigma en la búsqueda de nuevos fármacos

Hace más de dos décadas, el paradigma dominante en el campo del descubrimiento de fármacos proponía el desarrollo de ligandos con máxima selectividad que actúen modulando un único blanco molecular asociado a la enfermedad objetivo. En el marco de este paradigma, la selectividad y la potencia eran los aspectos esenciales para decidir si un candidato a fármaco progresaría a fases ulteriores de desarrollo.

Tal enfoque reduccionista se basa en dos premisas: a) fármacos altamente específicos evitarían efectos secundarios debidos a interacciones del mismo con biomoléculas distintas al blanco molecular, lo que conduciría a terapias más seguras y; b) al menos algunas enfermedades podrían ser tratadas adecuadamente modulando un único blanco molecular (paradigma "un gen, una droga, una enfermedad"). Esta aproximación se ajusta a la metáfora de que los fármacos ideales deberían actuar como ***balas mágicas*** (en inglés *magic bullets*), expresión que emplea por primera vez Paul Ehrlich (1854-1915) a principios del siglo pasado para referirse a los compuestos que en forma específica actuaban contra los microorganismos causantes de una enfermedad con una mínima toxicidad sobre el huésped.

Sin embargo, observaciones recientes han desafiado el paradigma anterior al detectarse que los fármacos utilizados actualmente en clínica a menudo se unen a más de un objetivo molecular. Bajo esta nueva perspectiva es que se sugiere que una analogía de ***escopeta mágica*** (*magic shotgun*) podría ser más adecuada para describir el mecanismo de acción de muchos fármacos conocidos, en particular dentro de ciertas categorías terapéuticas.

Algunos autores se han referido a este tipo de ligandos multiobjetivo como "drogas promiscuas". Por otro lado, si a un blanco molecular se unen diferentes ligandos químicamente heterogéneos, también se lo puede considerar como promiscuo. No obstante, preferiremos el uso de los términos ***ligando poliespecífico, multiobjetivo o multifuncional*** sobre el de promiscuidad, bajo la perspectiva de que los mismos evocan mejor la idea de una interacción selectiva con más de un blanco molecular o, si se quiere, una promiscuidad selectiva.

Las interacciones de los fármacos con proteínas distintas al blanco molecular tradicionalmente se han visto como un aspecto indeseable asociado a efectos secundarios no deseados. Pero en muchos casos —que van desde ciertos medicamentos psiquiátricos a las modernas terapias contra el cáncer— esta poliespecificidad está intrínsecamente vinculada a la eficacia terapéutica del fármaco.

Varios autores proponen que el diseño de fármacos multiobjetivo que interactúan con varios blancos moleculares asociados a una misma enfermedad conducirán a medicamentos nuevos y más eficaces para una variedad de trastornos de etiología compleja, entre ellos diversos desórdenes del sistema nervioso central, tales como enfermedades neurodegenerativas, desórdenes psiquiátricos y epilepsia.

Por otra parte, esta nueva visión es muy interesante ya que nos permite también concebir y explicar el ***reposicionamiento de fármacos*** o búsqueda de segundos usos médicos (que se describe más adelante en este capítulo) como una alternativa viable para desarrollar nuevos medicamentos. En este caso, pensaríamos en un fármaco multiobjetivo que interactúe selectivamente con blancos moleculares asociados a distintas enfermedades o desórdenes, o en utilizar

un fármaco (aprobado para una indicación) para una segunda indicación en la que esté involucrado el mismo blanco molecular o una isoforma del mismo blanco molecular.

1.5. Diseño y Búsqueda Racional de Fármacos

En el caso de la Química Medicinal las herramientas que se han constituido en máximos exponentes de racionalización en la búsqueda de nuevos fármacos son las relaciones cuantitativas estructura-actividad y el modelado molecular, aplicados ambos al diseño y búsqueda de nuevos agentes terapéuticos. Sin embargo, puede observarse en la ciencia moderna una tendencia a complementar la aproximación racionalista con aproximaciones menos racionales, legitimándose la incorporación de procesos cognitivos distintos del analítico-racionalista-reduccionista a la labor científica.

Hacia finales de la década de los años 80 se comenzó a trabajar sobre la idea del llamado “diseño racional” de nuevos fármacos, o diseño de fármacos guiado por computadoras (CGDD, de *computer-guided drug design*), entre otras denominaciones. La idea inicial, conceptualmente sencilla, era usar las tecnologías relacionadas con la biología estructural para generar la estructura tridimensional de una proteína diana (habitualmente por difracción de rayos X sobre la proteína cristalizada), y a partir de este conocimiento generar *in silico* pequeñas moléculas que pudieran servir de base para el desarrollo de fármacos. Habitualmente, una vez se tiene la estructura 3D, se utilizan programas quimioinformáticos para predecir la conformación y orientación de un ligando en el sitio de unión al receptor como por ejemplo el docking molecular (nos explayaremos sobre esta metodología en el Capítulo 6). De alguna forma, en origen se pretendía contraponer este paradigma al del screening sistemático, en el cual se confía en el azar para encontrar un punto de partida útil. Inicialmente este tipo de abordajes empezaron aplicándose en el marco de los procesos de optimización de líderes, pero posteriormente se han extendido también al descubrimiento de hits, mediante estrategias tales como el cribado virtual (Capítulo 6).

En resumen, nos referimos a metodologías de diseño racional de fármacos, cuando el objetivo es generar, crear, diseñar un fármaco nuevo (*de novo*). En los casos en los cuales empleamos las metodologías computacionales sobre bibliotecas virtuales de compuestos, que pueden existir o no, hablamos de “búsqueda racional”. En la búsqueda racional empleamos metodologías computacionales o in silico, tanto enfocándonos en el blanco molecular como en el ligando realizando un cribado virtual sobre dichas bibliotecas de compuestos (un detalle de las diferentes metodologías lo podrán encontrar en los Capítulos 5 y 6).

1.6. Reposicionamiento de Fármacos

El concepto de *reposicionamiento de fármacos* (en inglés *drug repurposing* o *drug reprofiling*) se refiere a encontrar nuevos usos terapéuticos para medicamentos ya existentes, incluyendo aquellos que se encuentran en el mercado farmacéutico, fármacos discontinuados, abandonados

y en desarrollo. Dentro de la comunidad internacional el interés en dicha estrategia ha crecido notablemente en los últimos años, y en países desarrollados como Estados Unidos y Reino Unido se han generado diversos programas públicos o público-privados para promoverla.

El reposicionamiento de fármacos está intrínsecamente ligado al concepto del uso “fuera de lo indicado” (*off-label*) de un medicamento, que se refiere a la prescripción de un medicamento por un médico (basado en evidencia científica o empírica) para indicaciones que aún no han sido evaluadas ni aprobadas por autoridades de la salud. El uso fuera de lo indicado es una práctica muy frecuente en ciertas ramas de la Medicina (por ejemplo, Psiquiatría y Pediatría).

La búsqueda de segundos usos médicos de fármacos ya conocidos posibilita introducir tratamientos innovadores en el mercado reduciendo considerablemente los tiempos (reducción de entre 3 y 12 años) y costos asociados al desarrollo de un medicamento novedoso. Los candidatos reposicionados han (como mínimo) sobrevivido previamente a estudios toxicológicos preclínicos (Figura 1.4), en muchos casos clínicos y poseen un perfil farmacocinético adecuado, ya caracterizado.

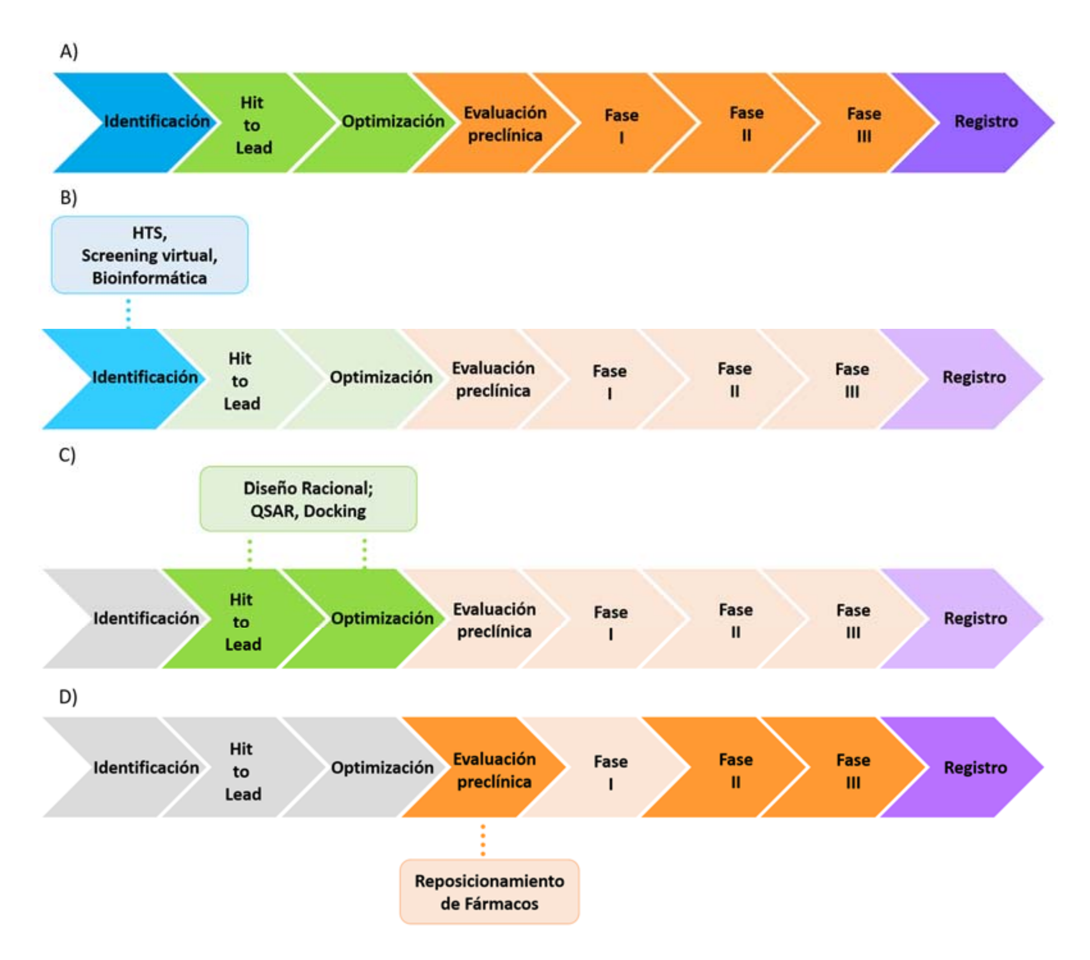


Figura 1.4: A) Proceso tradicional involucrado en el desarrollo de fármacos. Hit to Lead refiere al proceso de llegar a un compuesto líder partiendo de un conjunto de hits. B) La fase del proceso donde el screening de alta performance (HTS), el cribado virtual y la bioinformática están involucrados. C) Fases, donde aparecen las técnicas de optimización y comienza la etapa hit to lead. D) El fármaco reposicionado puede evitar la fase I en el proceso de descubrimiento del medicamento, solo si para el nuevo uso terapéutico es administrado en la misma forma farmacéutica y a dosis iguales o menores a las utilizadas en su uso original. (Adaptado de Bellera CL et al. *Curr Top Med Chem* 2018 ,18(5):369-381.

En cuanto al fármaco reposicionado, este ha sido previamente aprobado para su uso clínico y comercializado, y los posibles problemas de estabilidad y de manufactura ya han sido resueltos (por ejemplo, ya se han superado los estudios de pre-formulación). Finalmente, aquellos compuestos cuya patente se ha vencido pueden brindar soluciones relativamente poco costosas para nuevos problemas.

Los reposicionamientos exitosos probablemente han contribuido al interés de expansión de las indicaciones terapéuticas de fármacos ya conocidos. Por ejemplo, el Sildenafil originalmente fue investigado para el tratamiento de la hipertensión e isquemias cardíacas, pero se transformó en un éxito de ventas por su uso en el tratamiento de la disfunción eréctil. Por su parte la aspirina ha expandido sus indicaciones terapéuticas y hoy es ampliamente utilizada en la prevención de accidentes cerebro y cardiovasculares en pacientes con una condición cardiovascular preexistente. En la tabla 1.2 se observan diferentes ejemplos de reposicionamiento de fármacos.

Tabla 1.2: Ejemplos de Reposicionamiento de Fármacos

Fármaco	**Indicación Original**	**Nueva Indicación**
Aspirina	Inflamación	Antiplaquetario
Anfotericina B	Infecciones fúngicas	Leishmaniasis
Bromocriptina	Enfermedad de Parkinson, hiperprolactinemia y galactorrea	Diabetes mellitus
Celecoxib	Osteoartritis y artritis reumatoidea en adultos	Poliposis adenomatosa familiar, cáncer de colon y de mama
Finasteride	Hiperplasia prostática benigna	Pérdida de cabello
Fluoxetina	Depresión	Disforia Premenstrual
Metotrexato	Cáncer	Artritis reumatoidea, Psoriasis
Minoxidil	Hipertensión	Pérdida de cabello
Paclitaxel	Cáncer	Prevención de restenosis coronaria
Ropinirol	Hipertensión	Enfermedad de Parkinson y síndrome idiopático de las piernas inquietas
Sildenafil	Angina	Disfunción eréctil
Tadalafil	Inflamación y enfermedad cardiovascular	Disfunción eréctil
Topiramato	Epilepsia	Obesidad
Warfarina	Prevención de trombosis	Profilaxis secundaria seguida de infarto de miocardio
Zidovudina	Cáncer	HIV/SIDA

El reposicionamiento de fármacos asistido por computadora se basa en dos principios generales: a) Los medicamentos que comparten ciertas características moleculares mínimas, biológicamente relevantes, podrían interactuar con el mismo blanco molecular (enfoque centrado en los fármacos) y; b) los trastornos de salud relacionados con las mismas o similares proteínas desreguladas o disfuncionales pueden ser tratados con los mismos medicamentos (enfoque centrado en la enfermedad). Diversos métodos computacionales podrían ser útiles para revelar las relaciones proteína-fármaco o proteína-proteína ocultas. Una clasificación general en aproximaciones bioinformáticas, quimioinformáticas y basadas en análisis de literatura de alto rendimiento puede observarse en la figura 1.5.

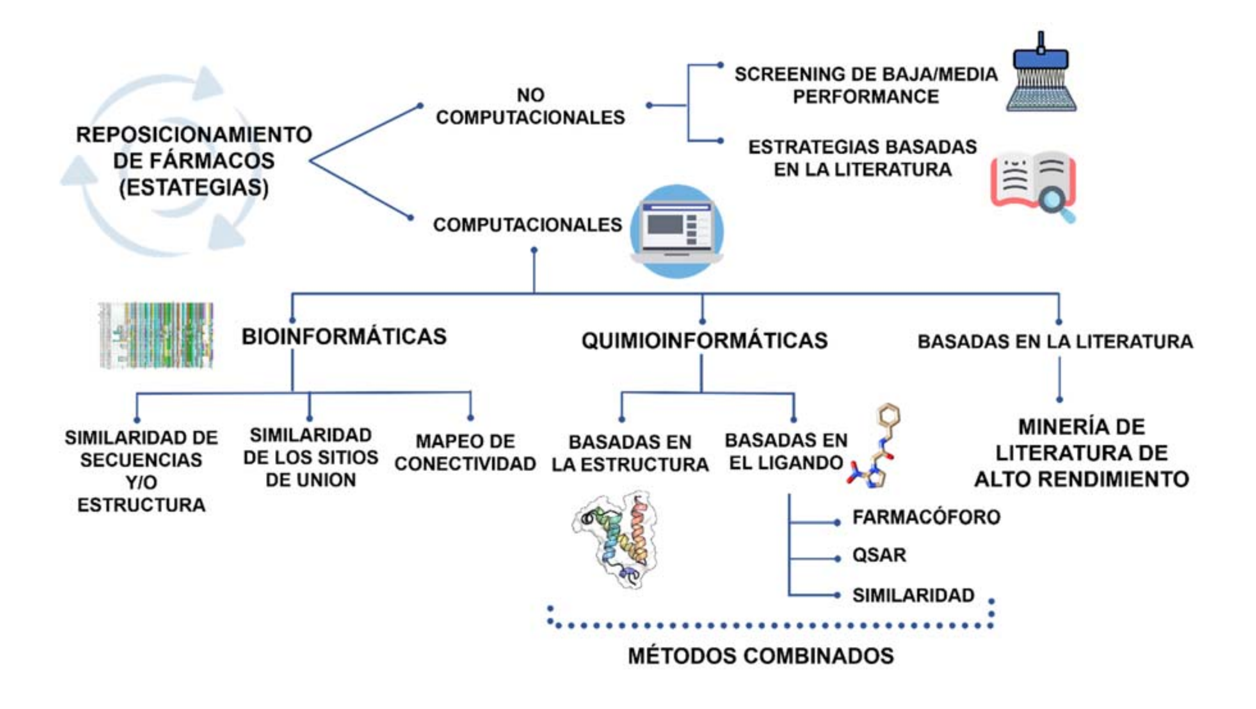


Figura 1.5: Metodologías aplicadas en Reposicionamiento de Fármacos Asistida por Computadora (Adaptado de Bellera CL et al Mini Rev Med Chem 2015).

Nota: en capítulos siguientes de este libro describiremos algunas metodologías que se emplean en el diseño, búsqueda y reposicionamiento asistido por computadora que estudiaremos en química medicinal. La integración de estas metodologías y otras más novedosas son de injerencia de la asignatura Nuevas estrategias aplicadas al descubrimiento de fármacos, asignatura electiva para estudiantes de la carrera de Farmacia de la UNLP.

Referencias

Bellera, C.L. et al (2015). High-throughput drug repositioning for the discovery of new treatments for Chagas disease. *Mini Review in Medicinal Chemistry*, 15(3), 182-193.

Bellera, C.L. et al. (2018). Modern Approaches for the Discovery of Anti-Infectious Drugs for the Treatment of Neglected Diseases. *Current Topics and Medicinal Chemistry*, 18(5), 369-381.

Cook, D. et al. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. *Nature Reviews Drug Discovery*, 13(6), 419–431.

Mohs, R.C. & Greig, N.H. (2017). Drug discovery and development: Role of basic biological research. *Alzheimer's and Dementia: Translational Research and Clinical Interventions*, 60, 3(4), 651–657.

Roberts R.A., et al. (2014). Reducing attrition in drug development: smart loading preclinical safety assessment. *Drug Discov Today*, 19, 341-347.

CAPÍTULO 2
Origen de los Fármacos

Mauricio E. Di Ianni y Carolina L. Bellera

Históricamente, el uso de sustancias químicas capaces de interferir positiva o negativamente ya sea sobre la salud humana o el medio ambiente se remonta al menos al año 1500 a.C según los registros del antiguo Egipto. Sin embargo, más allá de la prueba y error, el proceso mediante el cual el ser humano generó el conocimiento sobre los fármacos fue un misterio hasta la edad media, cuando los alquimistas eran los principales proveedores de salud. A partir del siglo XV d.C, con el renacimiento, la observación empezó a transformar la medicina y a clarificar el arte de curar. Hoy en día, la medicina moderna gira en torno al fármaco: somos capaces de comprender una patología a nivel molecular y a partir de allí razonar de qué manera intervenir para revertirla, como por ejemplo con la simple administración de una sustancia química. En este capítulo se abordarán las distintas metodologías clásicas capaces de encontrar nuevos prototipos activos, los cuales luego podrán transformarse en principios activos en caso de ser útil en salud humana, o de manera más general, en fármacos.

Para comenzar, retomemos la definición de ***compuesto líder*** citada en la introducción de este libro: "Compuesto químico que manifiesta una actividad determinada, NO superada por otros compuestos y cuyos cambios estructurales realizados hasta el momento no han mejorado su actividad ni disminuido su toxicidad", en otras palabras, el líder es la estructura más favorable en cuanto a la relación existente entre efecto farmacológico deseado y efectos farmacológicos adversos. Tal líder, surge de un subconjunto de estructuras que presentan un perfil farmacológico similar y que comparten un esqueleto estructural común al cual nos referiremos como ***prototipo activo***[1] o molécula prototipo. Sin embargo, el compuesto líder no siempre es la alternativa superadora, por el contrario, puede ser mejorado durante la última etapa del proceso de descubrimiento de fármacos: la optimización racional del líder, o simplemente ***optimización.***

En resumen, la generación de un nuevo candidato a fármaco conlleva dos etapas fundamentales:

1 - Descubrimiento del compuesto líder.
2 - Optimización racional del compuesto líder.

Cada una de estas etapas, involucra distintas metodologías las cuales se presentarán en detalle a continuación.

[1] No debe confundirse con el concepto de farmacóforo desarrollado más adelante en este libro.

1. Descubrimiento del compuesto líder

Como se mencionó anteriormente, el punto de partida consiste en encontrar un conjunto de moléculas que presenten una actividad farmacológica determinada. Por lo general, en esta etapa se observan moléculas con baja actividad y presencia de efectos adversos indeseados. Los mecanismos de generación del compuesto líder consisten fundamentalmente en la introducción de modificaciones químicas en la estructura de un prototipo activo y en el posterior testeo del impacto de estas modificaciones en la actividad biológica del producto final. Estas modificaciones pueden aumentar, disminuir o no influir en el perfil farmacológico del compuesto; lo que se busca idealmente es una gran potencia farmacológica (mucho efecto a dosis pequeñas) con la menor incidencia de efectos colaterales indeseados. A continuación, se detallan las distintas estrategias empleadas tanto para llegar a moléculas prototipo como a líderes.

1.1. Descubrimiento por casualidad (Azar, Serendipia)

Los hechos fortuitos o accidentales están presentes en todo el campo científico y la investigación farmacéutica no es la excepción. En relación a este hecho, el científico francés Louis Pasteur (1822-1895) enunció la famosa frase *“Dans les champs de l'observation le hasard ne favorise que les esprits préparés"* (En el campo de la investigación el azar no favorece más que a los espíritus preparados). A continuación, una serie de ejemplos ilustran la afirmación de Pasteur (Figura 2.1):

Durante la segunda guerra mundial, un barco estadounidense que trasportaba gas mostaza (mostaza sulfurada, una poderosa arma química) exploto en un puerto italiano. Se observó que muchos de los sobrevivientes que inhalaron el gas perdieron sus defensas naturales a cualquier tipo de microorganismo. Estudios posteriores mostraron que los afectados habían sufrido una destrucción masiva de sus glóbulos blancos. Es difícil visualizar como una droga que destruye células podría ser útil para la salud, sin embargo, este gas fue de gran importancia para el descubrimiento de agentes alquilantes (como las mostazas nitrogenadas) útiles en el tratamiento de diversos cánceres entre ellos la leucemia.

Los trabajadores de una fábrica de caucho manifestaban un rechazo inexplicable a cualquier tipo de bebidas alcohólicas. Esto se debía a la exposición prolongada de los obreros a un antioxidante utilizado en el proceso de elaboración del caucho, el cual evitaba la oxidación hepática normal del etanol conduciendo a la acumulación de acetaldheido en el organismo. Esto producía un estado desagradable al beber pequeñas cantidades de alcohol similar a una intoxicación aguda con alcohol, de manera que las personas evitaban beber. El antioxidante en cuestión, el disulfiram, se convirtió rápidamente en un medicamento utilizado para tratar el alcoholismo crónico.

Sildenafil: el caso ya es parte de la cultura popular. La droga fue inicialmente diseñada para tratar la hipertensión arterial y angina de pecho por su efecto vasodilatador. Durante los ensayos clínicos se observó una muy baja eficacia para estas patologías, y la aparición de

efectos colaterales llamativos. Actualmente la droga se comercializa exclusivamente para tratar la disfunción eréctil.

Gas Mostaza

Disulfiram

Sildenafil

Figura 2.1: Algunos casos de serendipia.

1.2. Medicina tradicional

Las civilizaciones antiguas dependían casi exclusivamente de su flora y fauna para poder sobrevivir. Como resultado de esto se fueron revelando los efectos de muchos frutos, hojas, raíces, flores, etc., con lo que surgieron muchos preparados con distintos usos terapéuticos a partir de distintos vegetales. Al mismo tiempo, el pase de generación en generación de este conocimiento fue incentivando la experimentación más profunda de muchas especies. Se citan a continuación, algunos ejemplos.

Dantrona: fármaco laxante diseñado a partir de las propiedades purgantes de las raíces de ruibarbo, efecto conocido hace varios siglos. Este fármaco está actualmente prohibido por la FDA por su alto potencial carcinogénico.

Artemisinina: fármaco antimalárico descubierto gracias a la investigación de la planta *Artemisia annua,* utilizada por los médicos chinos por más de dos mil años para el tratamiento de diversas dolencias, entre ellas la malaria.

Morfina: El uso del opio extraído de las capsulas de *papaver somniferum* se conoce hace más de 7000 años. Además de guerras, narcotráfico y adicción, el opio ha dado origen a una gran cantidad de medicamentos destinados principalmente a suprimir el dolor. A pesar de los numerosos efectos adversos que posee, ningún opioide moderno obtenido por síntesis total o semisintesis ha logrado superar el poderoso efecto analgésico del componente natural y mayoritario del opio: la morfina.

1.3. Aislamiento de productos naturales

La naturaleza es la fuente de compuestos biológicamente activos con mayor potencial de explotación y diversidad. La ciencia que estudia las drogas y principios activos de origen vegetal, animal y microbiano, es la ***farmacognosia*** y en ella se contemplan tanto extractos crudos como sustancias puras aisladas. La mayoría de los productos naturales biológicamente activos son metabolitos secundarios que presentan gran diversidad química y van desde moléculas pequeñas sencillas a moléculas muy grandes y complejas con gran cantidad de centros quirales. Estas características estructurales pueden observarse en los ejemplos de la figura 2.2.

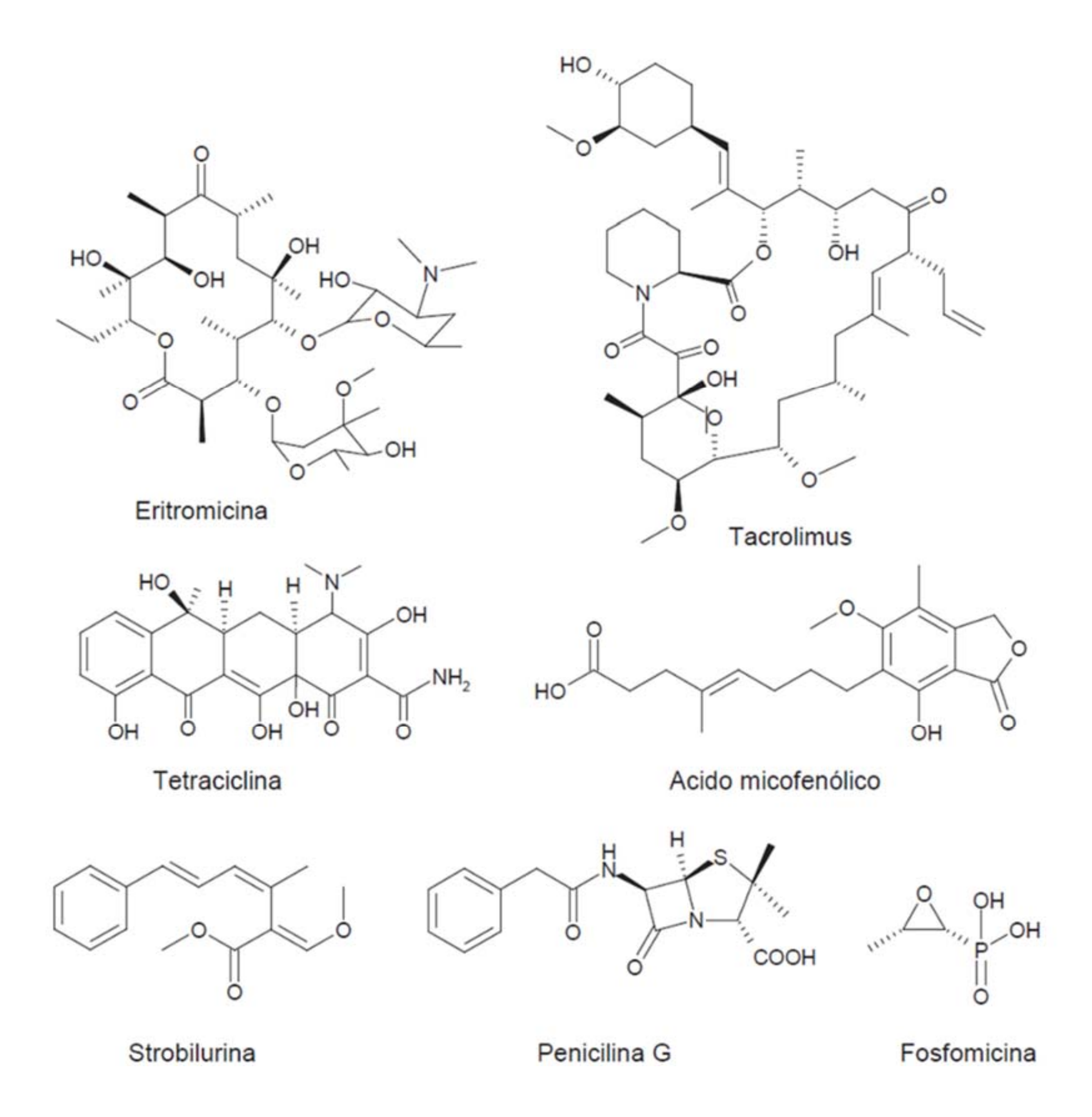


Figura 2.2: Ejemplos de productos naturales de utilidad terapéutica o en agroquímica, con diversos grados de complejidad estructural.

En la siguiente tabla se presentan diversos ejemplos de fármacos extraídos de distintas fuentes naturales.

Tabla 2.1: Ejemplos de fármacos obtenidos por diferentes fuentes naturales.

Fuente	Fármaco	Actividad Biológica
Vegetal	Morfina	Analgésico
	Cocaína	Anestésico local
	Quinina	Antimalarico
	Taxol	Anti cancerígeno
	Dogoxina	Glucósido Cardiotónico
Microorganismos (Hongos y Bacterias)	Penicilinas	Antibióticos
	Cefalosporinas	
	Tetraciclinas	
	Aminoglucósidos	
Animal	Tetrodotoxina	Veneno/ Toxina
	Batracotoxina	

1.4. Screening de alta performance (*High throughput screening*, HTS)

HTS es una metodología capaz de ejecutar una gran cantidad de ensayos de manera simultánea valiéndose de la robótica y tecnologías de avanzada. Para el caso de descubrimiento de fármacos, el HTS permite evaluar una gran cantidad de nuevos prototipos activos provenientes por ejemplo de síntesis orgánica, frente a una gran cantidad de ensayos biológicos distintos a fin de determinar qué actividad biológica poseen. Esto lo convierte en una metodología no racional de prueba y error que conlleva un costo económico considerable. La gran ventaja que posee es la velocidad y automatización del proceso. Se plica especialmente en aquellos casos en los que no se conocen los mecanismos farmacológicos responsables de la actividad de determinada droga.

1.5. Búsqueda de fármacos asistido por computadora (Cribado virtual)

Nuevos prototipos activos pueden encontrarse llevando a cabo análisis exhaustivos tanto de repositorios digitales de bibliografía científica o de bases de datos de estructuras químicas. Como se verá más adelante, estos estudios pueden conducirse en base al conocimiento del receptor molecular de interés (método directo) o en base al conocimiento de las características estructurales de las moléculas, responsables de determinada actividad biológica (métodos indirectos). En los Capítulos 5 y 6 nos dedicaremos exclusivamente a esta metodología.

1.6. Análisis de fármacos existentes (drogas “me too” y “me better”)

Muchas compañías farmacéuticas diseñan drogas basadas en moléculas ya existentes con el objetivo de competir en el mercado y eludir derechos de propiedad intelectual (patentes). Estos son los llamados fármacos “me too” (“yo también”). Son ejemplos de estos compuestos los anti-ácidos Ranitidina de Laboratorios Glaxo y su me too Famotidina de Laboratorios Merck. Otros ejemplos pueden observarse en la Figura 2.3 en relación con el Captopril y drogas “mee too” asociadas. Por otro lado, esta práctica en ocasiones conduce a una mejora en el perfil farmacológico de la droga, con lo que se habla de “me better drugs” (yo mejor).

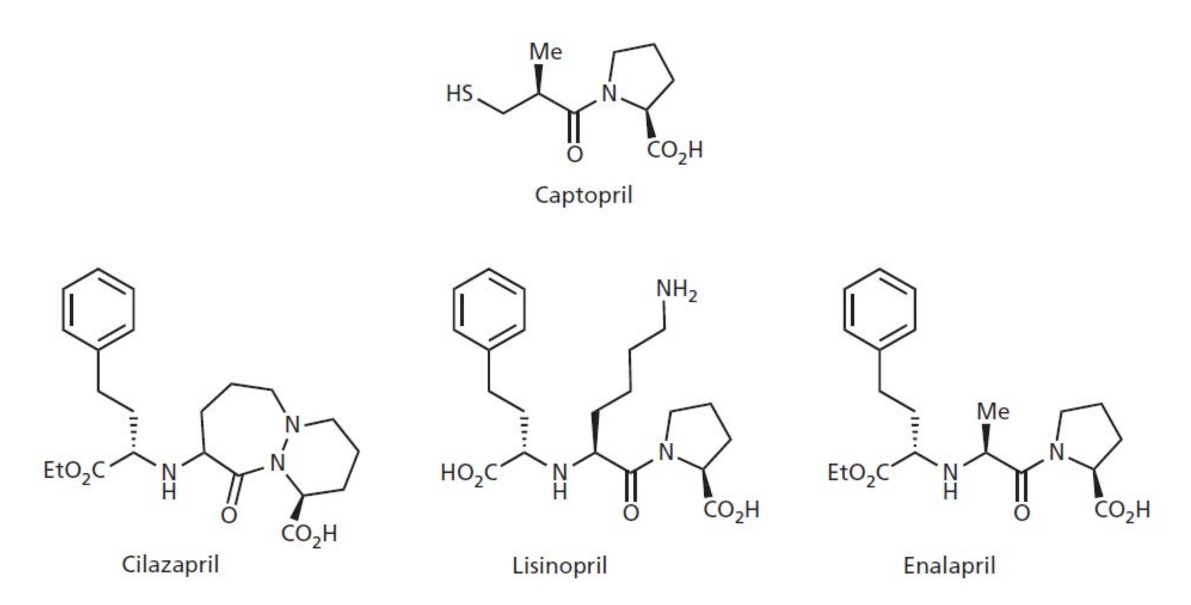


Figura 2.3: Captopril y drogas me too.

1.7 Aprovechamiento de efectos colaterales

Los efectos adversos de muchas sustancias puedes ser explotados como indicación terapéutica. Por ejemplo, dado que muchas sulfonamidas presentan efecto diurético a dosis altas, esta familia de compuestos que fue inicialmente desarrollada como antibacterianos, se comercializa también como fármacos diuréticos. Muchas veces, pequeños cambios en la estructura química, conlleva cambios significativos en la actividad farmacológica. De esta manera, una sulfonamida antibacteriana puede convertirse fácilmente en un diurético, un hipoglucemiante o un anticonvulsivo (Figura 2.4).

Hidroclorotiazida (Diuretico)

Zonisamida (Anticonvulsivo)

Sulfacetamida (Antibiotico)

Figura 2.4: Sulfamidas con distintas indicaciones.

2. Optimización del compuesto líder

La etapa final del diseño de fármacos es la optimización y es sin duda la más importante y ambiciosa. Entre los objetivos de esta etapa podemos mencionar los siguientes:

Selectividad: se espera que el fármaco interaccione únicamente con el blanco (o los blancos) para el cual fue diseñado.

Gran potencia: el fármaco debe tener un gran efecto a bajas dosis.

Ausencia de efectos adversos o efectos colaterales mínimos.

Propiedades farmacocinéticas aceptables: el fármaco debe ingresar satisfactoriamente al organismo, llegar al sitio de acción, surtir su efecto y ser eliminado dejando la menor cantidad de residuos o productos del metabolismo en circulación.

Toxicidad escasa o nula.

Estabilidad química

A continuación, se enumerarán las estrategias clásicas que apuntan a optimizar la interacción entre el fármaco y el blanco de interés responsable del efecto farmacológico.

2.1. Variación de sustituyentes

La variación química de sustituyentes (-R´s) fácilmente accesibles es uno de los métodos más comunes para ajustar la actividad de las moléculas.

Sustituyentes Alquílicos. Los más fáciles de alterar son los –R´s alquilo unidos a grupos amino, esteres y amidas. La técnica más extendida es la reacción de la molécula con un grupo dador del –R que mediante un ataque electrofílico o nucleofilico intercambia el –R original por el –R que se desea incorporar. La selectividad por un determinado sitio de acción puede conseguirse incorporando grupos alquílicos ramificados como los grupos isopropilo, isobutilo, etc. Un ejemplo conocido es del análogo de la Adrenalina, Isoprenalina, en el cual se consigue una selectividad por el receptor β-adrenérgico por sobre el -adrenérgico, mediante el intercambio de un metilo por un isopropilo (Figura 2.5).

Figura 2.5: Aumento de la selectividad de la Adrenalina.

Sustituyentes Aromáticos. Si una droga posee un anillo aromático, la posición de los sustituyentes que este posea puede variarse y el resultado puede ser un aumento de la actividad biológica. De la misma manera, el cambio de posición de un sustituyente impacta directamente sobre las características electrónicas y estéricas de sus vecinos. Por ejemplo, el efecto atractor de electrones de un grupo nitro afectará en mayor medida a la basicidad de un sustituyente amino, si este se encuentra en la posición –para por sobre la posición –meta.

2.2. Extensión de la Estructura

Esta estrategia implica la adición de grupos funcionales que aporten una vía de interacción extra o accesoria con el blanco a la molécula original. Esto es muy útil cuando, por ejemplo, el blanco molecular posee regiones en su sitio activo con distintas características estructurales, siendo una región hidrofílica, otra hidrofóbica, etc. De esta manera, pueden adicionarse sustituyentes con grupos polares para obtener un mayor número de puentes de hidrógenos o interacciones iónicas con el blanco molecular. Adicionalmente, con esta metodología es posible transformar un agonista en un antagonista en el caso de que la interacción conduzca a un estado inactivo del receptor o blanco molecular.

2.3. Extensión o contracción de cadenas laterales o anillos (homólogos, vinólogos)

Algunas drogas presentan dos importantes grupos de unión al blanco molecular separados por una cadena hidrocarbonada. En tal caso, es posible que la longitud de dicha cadena no sea la ideal para la mejor interacción. Por lo tanto, una táctica muy útil para modificar la actividad es acortando o alargando dicha separación. Si la cadena que separa los grupos activos está formada solo por sistemas de enlaces conjugados hablamos de vinólogos; en el caso de sistemas que involucran grupos metileno, se trata de series de homólogos.

2.4. Variación y Fusión de anillos

Esta estrategia es el caso emblemático de la controversia en torno a las drogas “me too”. Con el objetivo de evitar restricciones de propiedad intelectual, muchas compañías farmacéuticas reemplazan el anillo central (aromático o hetero-aromático) de un determinado fármaco, por otro similar (anillos fusionados, distinto heteroátomo o mayor número de carbonos en el anillo) consiguiendo estructuras activas similares, pero sin aportar mejoras significativas en el perfil farmacológico (misma potencia, mismos efectos adversos). Pese a ello, esta estrategia ha logrado dar con algunas estructuras que presentan ventajas significativas con respecto a las originales (drogas “me better”).

2.5. Isósteros

Desde el punto de vista clásico, los Isósteros son átomos o grupos de átomos que presentan la misma valencia y por ende poseen similaridades fisicoquímicas (Figura 2.6). Por ejemplo, -SH, $-NH_2$ y $-CH_3$ son isósteros de –OH. Mientras que –S, -NH y CH_2, son isósteros de –O. En un sentido más amplio. la generación de isósteros de una determinada molécula es una manera racional de modificación estructural, mediante la cual se pueden conocer los efectos del tamaño, polaridad, distribución electrónica y naturaleza de enlace sobre la actividad biológica. Por ejemplo, el Fluor puede considerarse un isóstero del Hidrógeno ya que poseen virtualmente el mismo tamaño. Sin embargo, este átomo es mucho más electronegativo que el hidrógeno, de manera que podemos conocer el efecto de la variación de las propiedades electrónicas en la actividad, sin una intervención significativa de los efectos estéricos.

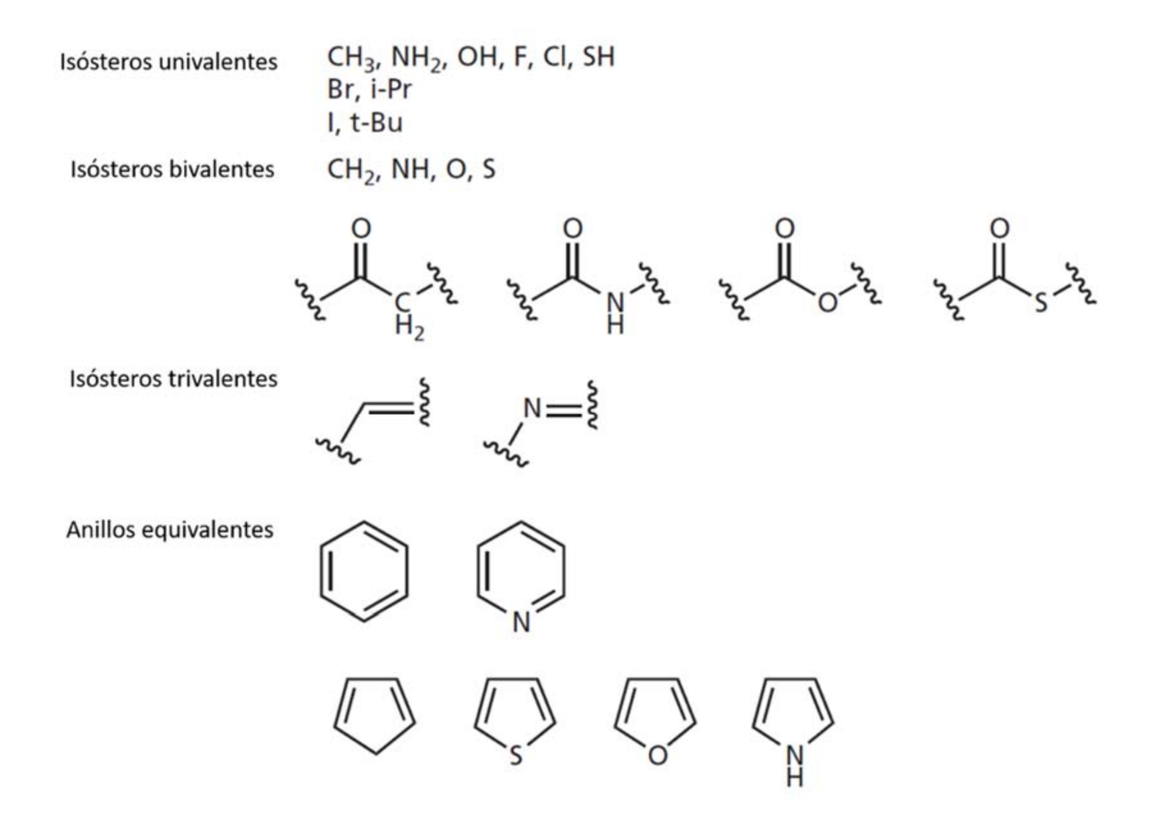


Figura 2.6: Ejemplos de isósteros clásicos.

2.6. Bioisósteros

Los bioisosteros son grupos que pueden utilizarse para reemplazar a otros, reteniendo la actividad biológica deseada. Son habitualmente utilizados para intercambiar grupos que son relevantes para la actividad biológica pero que por alguna razón generan problemas. Por ejemplo, los primeros antihistamínicos presentaban el grupo tiourea como el principal contribuyente al efecto farmacológico de esta familia de compuestos, pero que al mismo tiempo era el responsable de sus marcados efectos adversos. El remplazo de este grupo permitió conservar el efecto antihistamínico y eliminar los problemas de toxicidad (Figura 2.7).

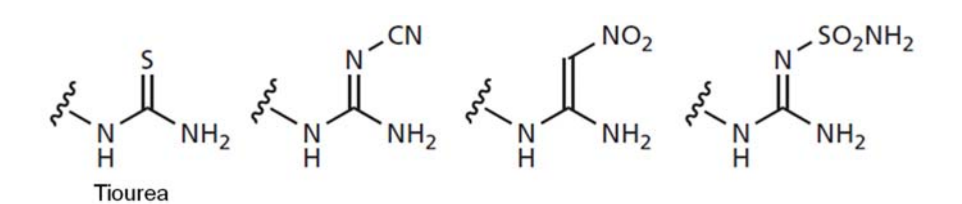


Figura 2.7: Bioisósteros de la tiourea.

2.7. Simplificación de la estructura

La simplificación molecular presenta grandes ventajas cuando se desean optimizar compuestos líderes provenientes de fuentes naturales, las cuales suelen ser de gran complejidad estructural. Una vez que la porción molecular activa es identificada, es posible reemplazar las partes no esenciales sin una perdida en la actividad biológica. Un ejemplo representativo de esta estrategia es la simplificación molecular de la cocaína, para dar con anestésicos locales más fáciles de obtener en el laboratorio mediante síntesis total (Figura 2.8).

Cocaína → Procaína

Figura 2.8: Simplificación molecular de la cocaína.

La gran ventaja de la simplificación molecular es el hecho de poder aprovechar las estructuras líderes provenientes de la naturaleza sintetizando derivados más sencillos pero que conservan la misma actividad farmacológica. Al mismo tiempo, una disminución moderada de la complejidad puede llevar a un aumento de la selectividad del fármaco. Sin embargo, una simplificación excesiva resulta en estructuras altamente inespecíficas (promiscuas) con diversos efectos terapéuticos, pero también marcados efectos adversos.

2.8. Rigidización de la estructura

Las moléculas flexibles, con varios enlaces capaces de rotar, pueden adquirir un gran número poses o formas en el espacio, las cuales se conocen como **confórmeros** o **conformaciónes**. Sin embargo, solo una de estas posiciones es la que satisface más eficientemente los requerimientos estructurales del blanco molecular para desencadenar la respuesta farmacológica. Esta pose se conoce como **conformación activa.** El acceso a un gran número de conformaciones en un compuesto flexible puede ser una desventaja, dado que es posible que este pueda adaptarse a otros sitios activos, dando origen a efectos indeseados. En el hipotético caso de que se conozcan los requerimientos estructurales para la interacción con el blanco molecular, la estrategia de rigidización molecular permitirá obtener una molécula capaz de activar dicho blanco de interés sin la aparición de efectos adversos.

En el ejemplo de la figura 2.8 se puede observar a la Diazepina (I), un inhibidor de la agregación plaquetaria. Este fármaco se une a su blanco molecular mediante el grupo guanidina y el sistema de anillos, los cuales se encuentran unidos entre sí por una cadena altamente flexible. Las estructuras (II) y (III) son ejemplos de compuestos activos en los cuales la cadena que conecta la guanidina y el biciclo ha sido modificada por la introducción de grupos funcionales rígidos.

Figura 2.9: Rigidización de cadenas flexibles.

Referencias

Alagarsamy, V. (2010) Introduction to Medicinal Chemistry. En: Textbook of Medicinal Chemistry, p1. Elsevier.

Graham L. P. (1995) An Introduction to Medicinal Chemistry. 5th ed. Oxford University Press, New York. ISBN 0-19-855872-4.

Wermuth C. et al (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1997). *Annual Reports in Medicinal Chemistry*, 33, 385–395.

CAPÍTULO 3
Síntesis de Fármacos

Laureano L. Sabatier, María L. Villalba y Luciana Gavernet

1. Rol de la síntesis química en el proceso de desarrollo de un fármaco

La mayoría de los fármacos son moléculas relativamente pequeñas que han sido pensadas y preparadas por químicos medicinales con experiencia en síntesis orgánica. Antes de comenzar una síntesis esta debe planificarse, es decir, plantear las rutas sintéticas que permitan la obtención de la molécula objetivo.

Denominaremos síntesis total a la síntesis química de un compuesto, a partir de precursores relativamente sencillos y disponibles comercialmente. Las materias primas en este tipo de síntesis habitualmente son compuestos derivados del petróleo de estructura simple. Por ejemplo, el Citalopram es un inhibidor selectivo de la recaptación de serotonina empleado para el tratamiento de la depresión, ataques de pánico, fobia social, etc. El enantiómero *S* del racemato Citalopram se denomina Escitalopram, y actualmente ambos fármacos están disponibles comercialmente. En el esquema 3.1 se representa la síntesis total del Escitalopram, en donde se puede apreciar que el material de partida es la ftalimida, un compuesto orgánico sencillo y asequible.

Esquema 3.1: Síntesis de Escitalopram.

En una síntesis parcial o semisíntesis, se parte de un producto natural que no ha sido previamente sintetizado, sino extraído y purificado de organismos por métodos de separación de mezclas. Esteroides y antibióticos β-lactamicos son casos de compuestos activos que se obtienen por semisíntesis. Otro ejemplo es la Heroína, un opioide con propiedades analgésicas y habitualmente empleando como droga recreativa. Tal como se puede apreciar en el esquema 3.2, su síntesis consiste simplemente de la acetilación de la morfina (un producto natural extraído de Papaver somniferum).

Esquema 3.2: Semisíntesis de Heroína.

2. Síntesis lineal y convergente

Cuando el objetivo de la síntesis es obtener un compuesto estructuralmente sencillo, se suele emplear una serie de pasos que son llevados a cabo uno tras otro hasta obtener el producto deseado. Este tipo de síntesis en donde el compuesto objetivo es el resultado de una serie de transformaciones lineales, se lo denomina síntesis lineal; y a los compuestos químicos producidos en cada etapa se les denomina intermedios sintéticos. Por ejemplo, la síntesis del Fentanilo (Agonista opioide) es de tipo lineal (Esquema 3.3).

Esquema 3.3: Síntesis lineal de Fentanilo.

Para moléculas más complejas, una síntesis convergente es con frecuencia preferible. En este caso se sintetizan dos o más fragmentos separadamente que luego al unirse generan el compuesto deseado. Las síntesis convergentes presentan las ventajas de ser más cortas, fáciles de llevar a cabo y generan rendimientos globales más elevados. Un ejemplo es la síntesis de Haloperidol (un antipsicótico convencional) que se representa en el esquema 3.4.

Esquema 3.4. Síntesis convergente de Haloperidol.

Debe mencionarse también que, además del rendimiento, resulta importante medir la pureza de los compuestos sintetizados. Definiremos como s*ustancia pura* a aquella que, sometida a distintos procesos de purificación, mantiene sus propiedades fisicoquímicas invariantes. Además, en el caso de un fármaco es particularmente relevante el estudio de la identidad de las

impurezas: no es lo mismo decir que un compuesto tiene un 5 % de agua como impureza que un 5% de cianuro, sales, solventes residuales, etc.

3. Planificación de Síntesis

En relación con la planificación de una síntesis química, estas metodologías han cambiado en el tiempo y se han optimizado con el desarrollo de la química. Una estrategia es la *asociación directa*, utilizada a en siglo XIX y a principios del siglo XX para la preparación de moléculas objetivo. Consiste en reconocer dentro de la estructura un número de subestructuras o unidades accesibles, las cuales se unen usando reacciones conocidas y preferentemente en un solo paso. Un ejemplo es la síntesis de colorantes azoicos del Esquema 3.5.

Esquema 3.5: Síntesis de colorantes azoicos mediante asociación directa.

Otra estrategia más general y muy utilizada se denomina *análisis retrosintético*, diseñada por Corey en la década del 60. Esta metodología permite postular en forma teórica diferentes rutas sintéticas para la molécula objetivo, mediante la desconexión de sus enlaces. La condición para que esta desconexión sea útil es que exista una reacción química capaz de reconectar los fragmentos generados. Además, deben priorizarse fragmentos relativamente estables y lograr la mayor simplificación posible de las reacciones.

Se definen como *sintones* a aquellos fragmentos estructurales (usualmente de carácter teórico) que son producto de una desconexión. Por otra parte, aquel compuesto orgánico (o reactivo) de existencia real que se utiliza en la práctica como fuente de generación del sintón se denomina *equivalente sintético*. Dentro de las reacciones de síntesis más comunes están las del tipo polar, las cuales pueden ser analizadas mediante una ruptura heterolítica teórica que genera un sintón que posee un átomo polarizado negativamente (un donador de electrones) y otro que

posee un átomo polarizado positivamente (un aceptor de electrones). Por ejemplo, en la síntesis de un éster (Esquema 3.6) se puede desconectar el enlace C-O para generar un sintón catiónico y un aniónico. Nótese que las retrosíntesis se representan con una flecha de doble trazo pues son reacciones antitéticas, de modo de diferenciarlas de las reacciones opuestas, las sintéticas, que utilizan flechas de un solo trazo.

Esquema 3.6: Posible desconexión de enlace de un éster para la generación de sintones.

El paso siguiente es la propuesta de equivalentes sintéticos. En el caso del ejemplo existen varias estrategias sintéticas como por ejemplo la reacción de fenol con el ácido, o con el anhídrido, o el cloruro de ácido o el éster correspondiente (entre otras reacciones). Es decir, que un sintón puede tener asociado más de un equivalente sintético (Esquema 3.7).

Esquema 3.7: Posibles equivalentes sintéticos para la síntesis de un éster.

En el caso de la Benzocaína (un anestésico local) el análisis retrosintético puede plantear como etapa inicial también la ruptura de la unión éster (Esquema 3.8). Como equivalentes sintéticos podemos elegir, por ejemplo, el ácido p-aminobenzoico y el etanol como se vio en el ejemplo anterior. Posteriormente puede proponerse la desconexión de los grupos funcionales de dicho acido, sin embargo, no existirá un equivalente sintético que mediante una reacción simple lo

regenere, ya sea desde la anilina o desde el ácido benzoico. Debe recurrirse entonces a una estrategia llamada interconversión del grupo funcional mediante la introducción de funciones latentes, esto es, grupos que están presentes en los equivalentes sintéticos pero que no son parte de la molécula objetivo.

Esquema 3.8: Análisis retrosintético de Benzocaína.

La función amina se puede generar a partir del grupo nitro mediante reacciones de reducción conocidas, por lo que este último grupo puede considerarse como función latente. Similarmente, el grupo carboxilo puede generarse a partir de la oxidación de un grupo metilo unido a un anillo aromático mediante la reacción de oxidación de cadena lateral (Esquema 3.9). Estas funciones latentes tienen la ventaja de poder introducirse al anillo de benceno fácilmente mediante reacciones de sustitución electrofílica aromática (nitración en medio acido o alquilación de Friedel-Crafts, respectivamente).

Esquema 3.9: Planteo retrosintético mediante el uso de funciones latentes.

La siguiente etapa sería la desconexión del p-nitrotolueno a tolueno. Este este último es un reactivo accesible y posee un grupo metilo dador de electrones (y orientador en las posiciones orto y para preferentemente), por lo que resulta más conveniente que la desconexión a nitrobenceno. Por lo tanto, la propuesta de síntesis de Benzocaína se muestra en el Esquema 3.10.

Esquema 3.10: Propuesta de síntesis de Benzocaína de acuerdo con el análisis retrosintético.

4. Síntesis de fármacos utilizando química combinatoria

Como se ha mencionado previamente, el proceso de descubrimiento de fármacos es costoso tanto en términos de tiempo como de dinero invertido. Sin embargo, han ocurrido avances como la automatización y miniaturización de ensayos biológicos que han producido importantes ventajas por cuanto es posible hacer un cribado mediante métodos de alta eficiencia (en inglés, *High Throughput Screening*) de una gran cantidad de compuestos en forma simultánea. En este contexto, la etapa limitante para el estudio de la actividad biológica es la capacidad de sintetizar nuevas moléculas.

La *química combinatoria* es un conjunto de procedimientos que posibilitan la generación eficiente, simultánea y rápida de una gran cantidad de estructuras. Estas sustancias diferentes se denominan colectivamente *bibliotecas de compuestos o quimiotecas*. A diferencia de la síntesis clásica, que busca obtener un solo compuesto bien caracterizado e identificado, con estas metodologías se espera preparar grandes quimiotecas que posteriormente puedan ser sometidas a ensayos biológicos. Por lo tanto, la química combinatoria resulta útil como fuente de nuevos compuestos líderes y también como herramienta para la preparación de varios miembros de una misma familia estructural para la optimización de fármacos.

La estrategia utilizada por la química combinatoria es tomar un conjunto de compuestos diferentes, pero químicamente relacionados entre sí, para que reaccionen con otro grupo de estructuras (también relacionadas entre ellas), generando bajo las mismas condiciones de reacción todos los productos posibles surgidos de la combinación de ambos grupos constructores (Figura 3.1).

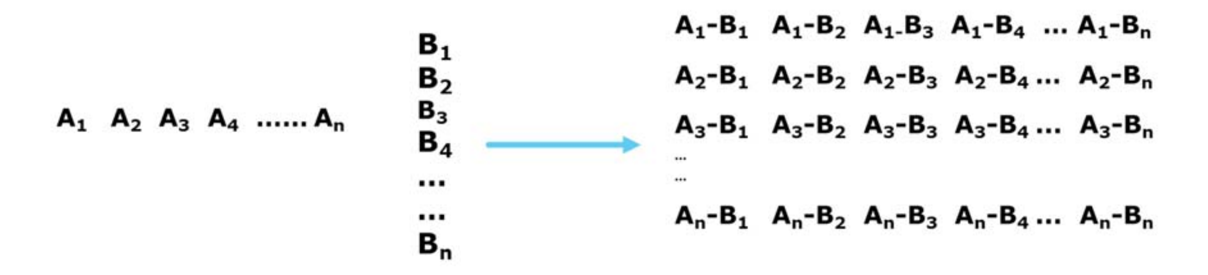


Figura 3.1: Representación de la estrategia de síntesis mediante química combinatoria.

La simultaneidad de las reacciones puede darse de dos maneras: mediante la síntesis de mezclas o por síntesis en paralelo (Figura 3.2). En el primer caso se obtienen distintos productos dentro de la misma mezcla de reacción, mientras que en el segundo caso cada compuesto se prepara en forma separada en un recipiente. En ambos casos se recurre a la síntesis en fase sólida como metodología de síntesis, dado que tiene como ventajas facilitar el proceso de automatización, y simplificar las etapas de aislamiento y purificación de los productos (las cuales usualmente representan un proceso largo y laborioso).

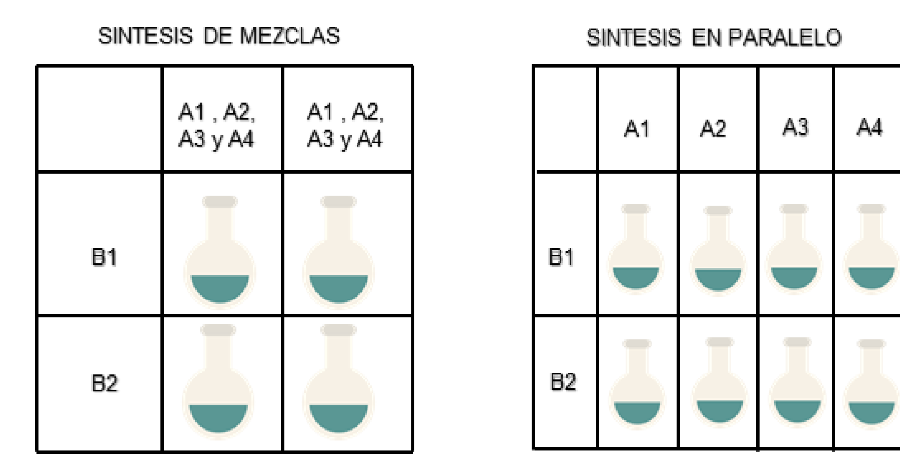


Figura 3.2; Representación de síntesis de mezclas y en paralelo. En el primer caso en cada recipiente se mezcla un reactivo B_i y varios A_i mientras que en el segundo se combina un reactivo B_i con solo uno del tipo Ai.

4.1 Síntesis en fase sólida

En la síntesis de fase sólida uno de los reactivos se une a un soporte sólido polimérico e insoluble en el solvente de reacción (denominado resina). Luego este compuesto "anclado" se hace reaccionar con el resto de los reactivos, quedando el producto final unido a la resina. Posteriormente un simple proceso de filtración elimina todo lo que se encuentra en solución (Figura 3.3). Esto permite por ejemplo trabajar con reactivos en exceso dado que aquellos

materiales que no reaccionan serán parte del filtrado en el proceso de aislamiento. La etapa final es la separación (o desacople) del producto deseado de la resina mediante el uso de algún reactivo adecuado. Este producto puede ser sometido posteriormente a procesos de purificación tradicional de ser necesario.

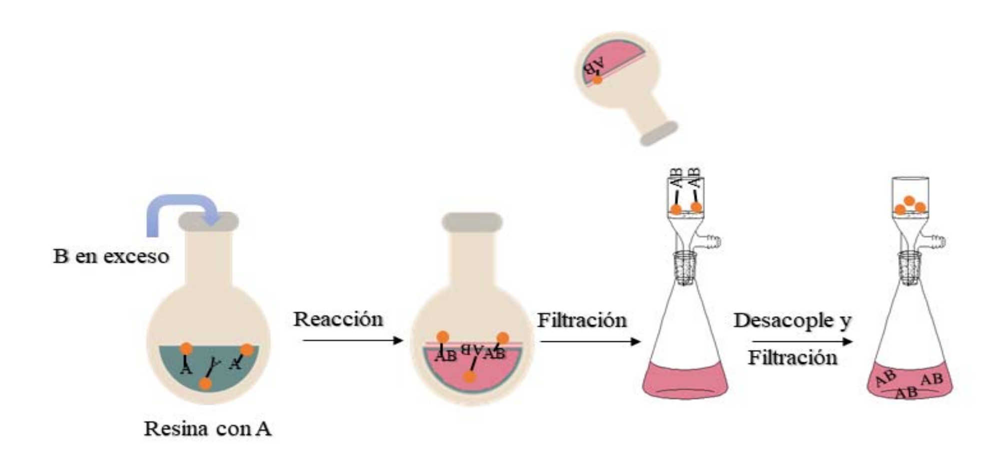


Figura 3.3: Esquema de síntesis en fase sólida. La resina se representa como una esfera naranja.

El empleo de este tipo de reacciones está basado en el desarrollo de la síntesis de péptidos de Merrifield en 1963 (Figura 3.4). Inicialmente el soporte sólido inerte está unido covalentemente a un grupo ligante, esto es, a una estructura que mediante una reacción química permita la unión del primer reactivo a la resina. En el caso de la síntesis de Merrifield los bloques constructores del producto final serán aminoácidos, por lo que estos reactivos presentan diferentes grupos funcionales (amino, ácidos carboxílicos, alcoholes, etc.). Por lo tanto, aquellas funciones que no deben reaccionar con el sistema ligante-resina deben encontrarse protegidas al momento de realizar el acoplamiento. Una vez realizado el acoplamiento inicial por medio del grupo ácido del aminoácido, se eliminarán por filtración el exceso de reactivos que no se han unido a la resina. Para eliminar residuos adsorbidos u ocluidos en la resina también se realizan lavados luego de la filtración. Habiendo inmovilizado el primer aminoácido se produce el crecimiento del péptido: se realiza la reacción de remoción del grupo protector del grupo amino para permitir la reacción con el siguiente aminoácido protegido. Ahora se tiene un dipéptido anclado a la resina que se aísla de los reactivos en exceso por filtración y lavado. Luego de ciclos repetitivos de acoplamiento-lavado-desprotección-lavado se obtiene el producto deseado, el cual se libera del soporte sólido, usualmente con un reactivo que permita remover también otros grupos protectores existentes en el péptido.

Del análisis de este ejemplo podemos concluir que para la síntesis en fase sólida es necesaria la presencia de un soporte polimérico inerte, entrecruzado e insoluble; un ligante capaz de conectar a dicho soporte el bloque constructor de la molécula objetivo y un conjunto de reacciones químicas que permitan la preparación del producto final, incluyendo reacciones de protección y desprotección de grupos reactivos.

En relación con el soporte insoluble, estos tienen la función de facilitar el aislamiento de los productos mediante separación por filtración. Sin embargo, los solventes penetran dentro de la resina en distintas proporciones expandiendo su volumen (fenómeno conocido como hinchado de la resina). Las resinas más usadas son granos de poliestireno entrecruzado con divinilbenceno, las cuales presentan características hidrofóbicas. En los anillos aromáticos de poliestireno se ubican, en ciertos intervalos, los grupos ligantes. Estos grupos deben formar enlaces estables con los reactivos utilizados para anclar los bloques constructores, pero deben ser capaces de escindirse cuando la reacción se completa para liberar el producto final. Dado que los grupos ligantes se encuentran distribuidos en toda la cadena polimérica del soporte, resulta importante también la selección de un solvente que pueda hinchar lo suficiente la resina para permitir la llegada de los grupos ligantes ubicados en el interior del polímero. Las resinas de poliestireno/divinilbenceno se hinchan apropiadamente con solventes apolares, pero su hinchamiento es pobre en solventes polares próticos.

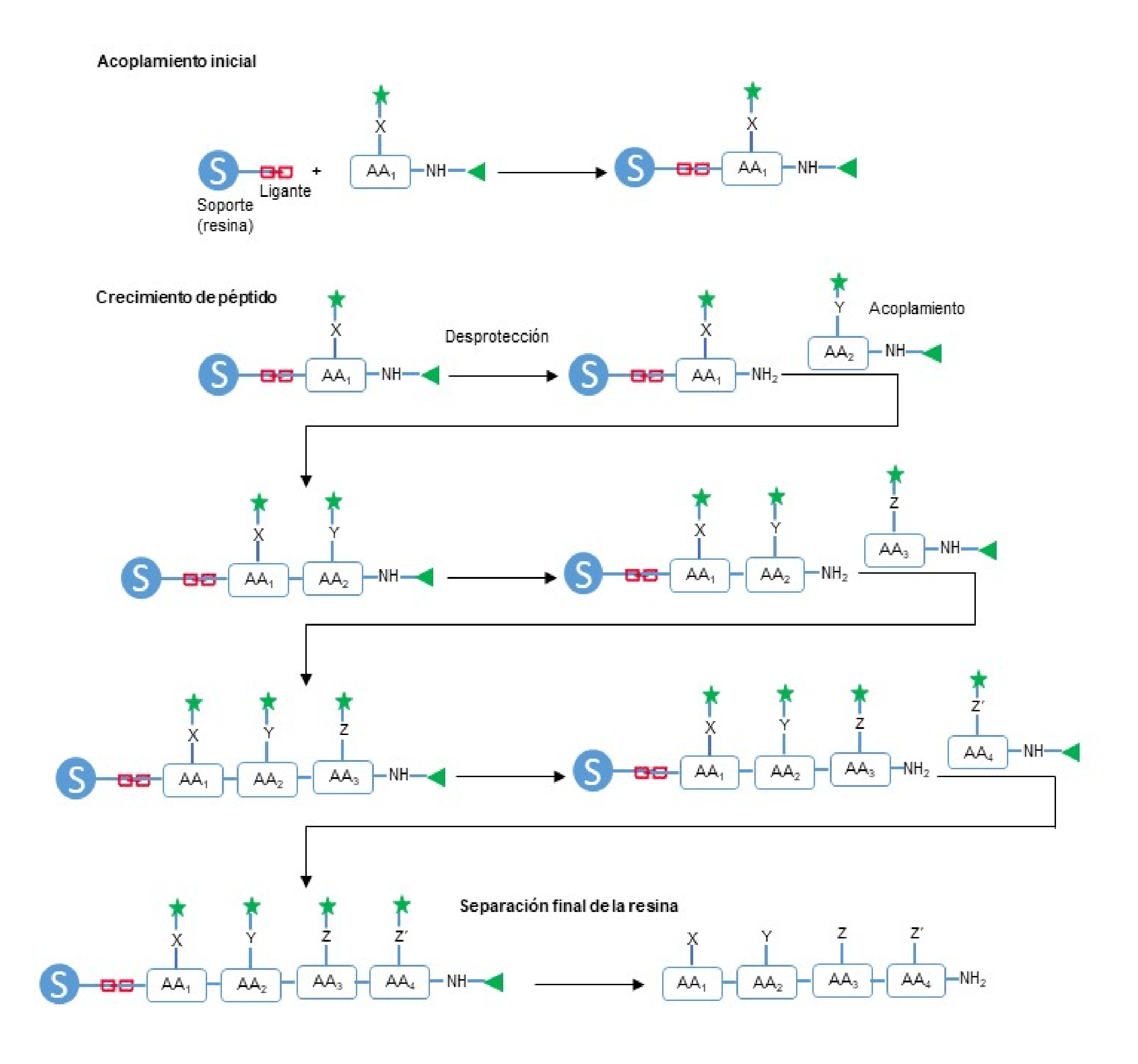


Figura 3.4: Representación de la síntesis en fase sólida de péptidos. Los grupos protectores se representan con formas de color verde.

En la Figura 3.5 se muestran las estructuras de algunos grupos ligantes utilizados frecuentemente en la síntesis de fase sólida. La resina de Merrifield es simplemente una derivatización de la cadena polimérica con el grupo clorometilo. Alternativas a esta fase sólida son la resina de Wang y la de Sasrin, las cuales poseen al menos un grupo alcoxi (dador de electrones) en el anillo aromático, lo que las hace más lábiles a ácidos que la resina de Merrifield. Por el contrario, la resina PAM es más resistente a ácidos ya que posee un ligante con un grupo atractor de electrones. Otros ligantes como las resinas de Rink amida, permiten la preparación de amidas primarias y las resinas dihidropirano presentan alternativas de anclaje, ya que en vez de grupos ácidos pueden incorporarse en este caso grupos alcoholes por adición al doble enlace. Finalmente pueden mencionarse las resinas fotolábiles, las cuales permiten el desacople del producto final por irradiación con luz ultravioleta.

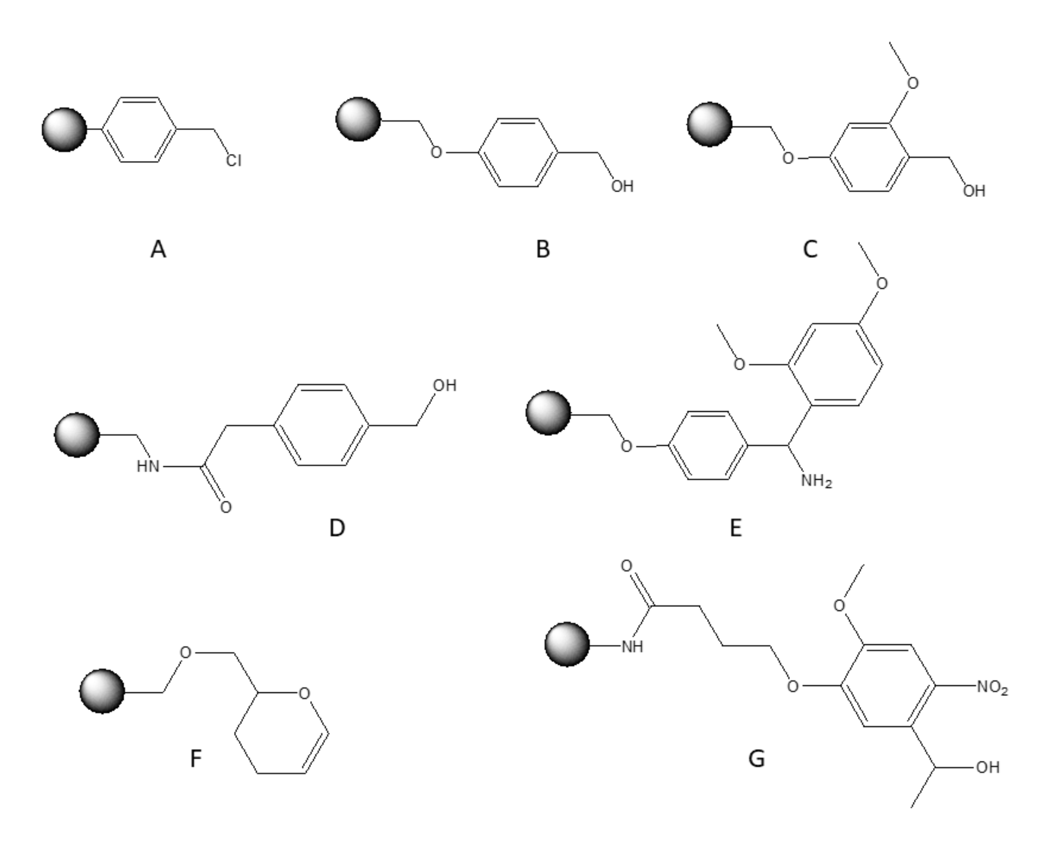


Figura 3.5. Estructura de grupos ligantes unidos a resinas utilizadas en síntesis en fase sólida. A: resina de Merrifield; B: resina de Wang; C: resina de Sasrin; D: resina PAM, E: resina de Rink; F: resina de dihidropirano; G: resina fotolábil.

La síntesis en fase sólida representa una metodología con ciertas ventajas respecto de la síntesis tradicional en solución. Principalmente la presencia de un soporte sólido permite el trabajo con reactivos en exceso sin generar problemas en la etapa de aislamiento, dado que estos compuestos, junto con otros subproductos solubles, son fácilmente removidos. En este mismo sentido resulta muy útil para realizar reacciones en más de una etapa, ya que los productos de las reacciones intermedias no necesitan purificarse. En cuanto a las condiciones de reacción, se admiten solventes de alto punto

de ebullición, que luego no deben evaporarse en el aislamiento como en la síntesis tradicional. Además, puede trabajarse con una mayor seguridad con sustancias tóxicas ya que al estar unidas al soporte sólido son fácilmente manipulables.

Por otra parte, las resinas pueden ser regeneradas y reutilizarse; y en general las reacciones son más amigables con el medio ambiente ya que prescinden de solventes y otros materiales que usualmente se utilizan en las etapas de aislamiento en la síntesis tradicional. Adicionalmente mediante este tipo de soportes se facilitan reacciones químicas que resultan dificultosas en la síntesis tradicional. Por ejemplo, se favorecen las reacciones intramoleculares (como las ciclaciones) debido a las condiciones de pseudodilución cercanas a los sitios de reactivos de la resina (Esquema 3.11).

Esquema 3.11. Síntesis de hidantoínas en fase sólida. Las esferas rojas representan el soporte sólido.

Además, la síntesis en fase sólida puede automatizarse para la generación de bibliotecas combinatorias ya sea mediante técnicas de síntesis de mezclas o en paralelo.

4.2 Bibliotecas de compuestos en mezclas

Dentro de los métodos de preparación de bibliotecas en mezclas, el más utilizado es la estrategia de mezclar y separar, desarrollado por Furka en 1988. Esta metodología se ejemplifica con la síntesis de tripéptidos que contienen leucina (Leu), alanina (Ala) y/o valina (Val) como bloques constructores (Figura 3.6). Inicialmente los granos de resina de dividen en partes iguales y se los someten a la reacción con el primer aminoácido, el cual es diferente en cada caso. Luego de que la reacción se completa se filtran, lavan y mezclan las resinas para finalmente dividirlas en 3 fracciones. Estas porciones contienen a los aminoácidos unidos a las resinas, pero no hay resina que contenga más de un tipo de aminoácido acoplado, dado que los reactivos se agregaron puros. Luego estos 3 grupos reaccionan nuevamente con los 3 aminoácidos en forma separada, obteniéndose ahora dipéptidos. Nuevamente se filtran, lavan y mezclan, para generar un conjunto de 9 dipéptidos unidos a las resinas.

El proceso puede repetirse las veces necesarias, dependiendo de la longitud del péptido deseado. En el caso del ejemplo, se obtienen 3 mezclas que en total generan 27 tripéptidos. Una vez sintetizadas las mezclas se procede a la evaluación biológica. Usualmente el testeo biológico implica la liberación de los compuestos de la fase sólida, aunque existen algunos ensayos que son compatibles con los compuestos unidos a la resina. Conocida la mezcla más activa debe determinarse cuál de las estructuras que la componen es la mayor responsable de la actividad. Este proceso se conoce con el nombre de *deconvolución de mezclas*.

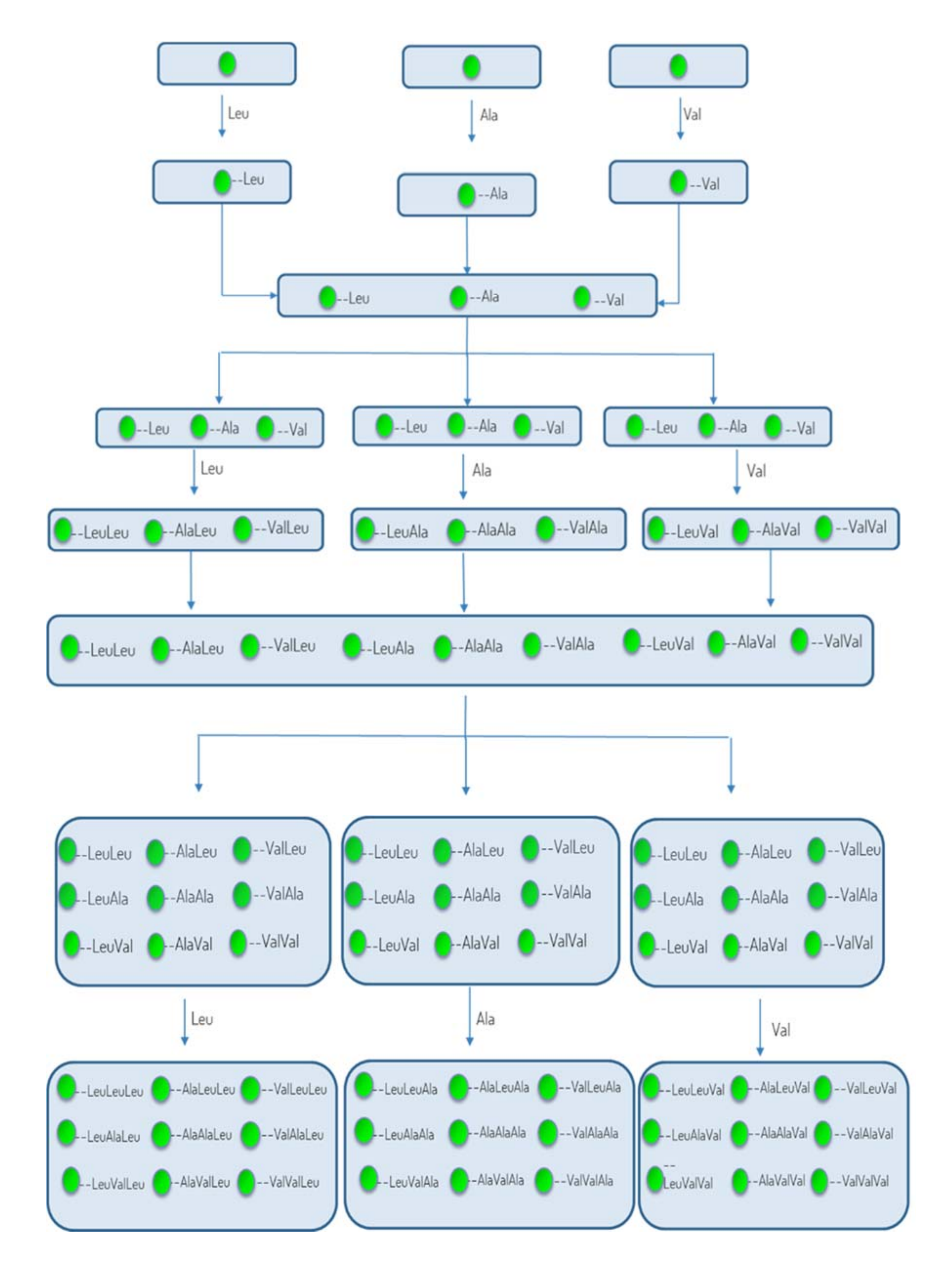


Figura 3.6: Estrategia de mezclar y separar para el desarrollo de síntesis combinatoria en mezclas.

En el caso del ejemplo, supongamos que la segunda mezcla es la más activa (Figura 3.7). Dado que se conoce el último reactivo agregado, se sabe la identidad del último aminoácido del péptido más activo, en nuestro caso Ala. Entonces se hace la primera resíntesis: se toman las

mezclas de dipéptidos (reservadas de la etapa anterior) y a todas se le agregan Ala. Se obtienen mezclas de tripéptidos que tienen Ala en la última posición y se ensayan. Según el ejemplo de la figura 3.7 la mezcla más activa es la primera, la cual, de acuerdo a la ruta sintética (Figura 3.6), tiene Leu como segundo aminoácido. Entonces se hace una nueva resíntesis que será ya de los aminoácidos unidos a la resina, y si el recipiente más activo es al primero puede deducirse que el candidato activo tiene estructura Leu-Leu-Ala.

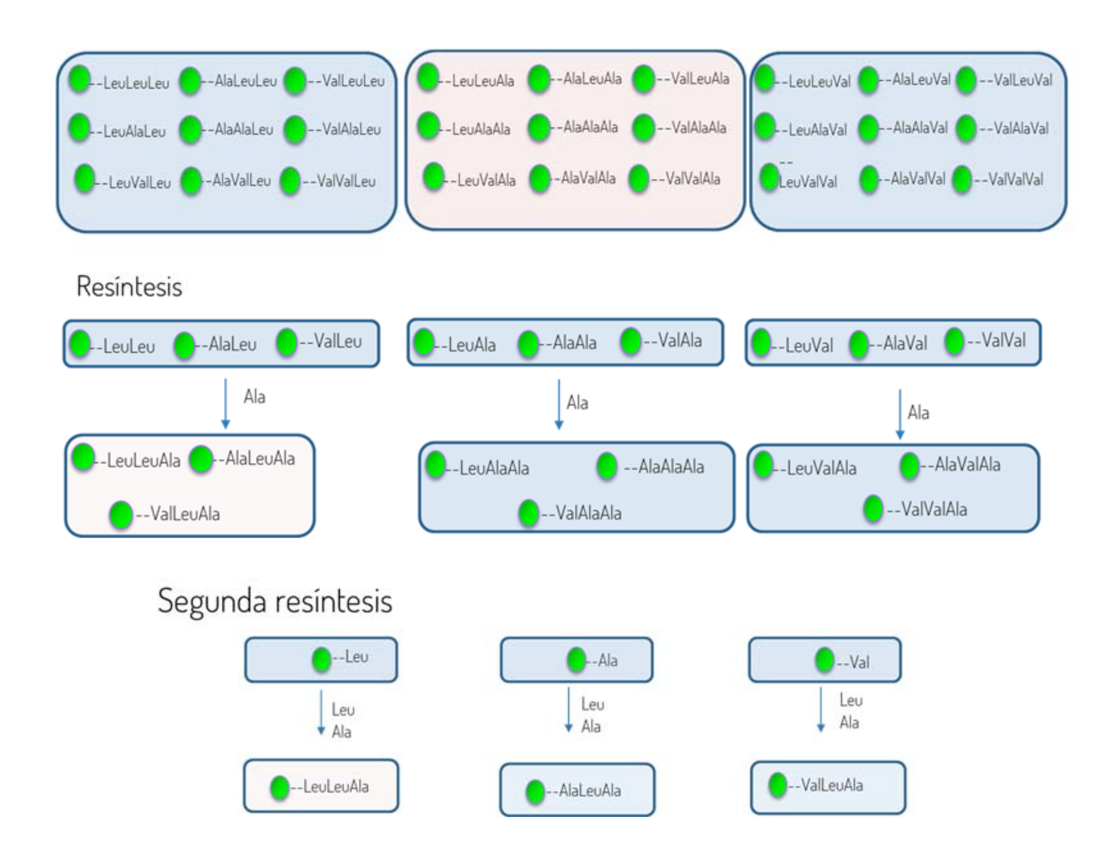


Figura 3.7: Proceso de deconvolución en la técnica de mezclar y separar. Las mezclas más activas se marcan en color rosa.

El método de mezclar y separar es útil para la preparación de gran cantidad de compuestos por combinación de sus bloques constructores, y posee una manera simple de identificar los compuestos activos por deconvolución. Sin embargo, debe tenerse en cuenta que es posible que las reacciones no generen los productos esperados, por lo que erróneamente puede concluirse que un compuesto es inactivo cuando en realidad este no se encuentra presente en la mezcla. Adicionalmente, es posible que un compuesto activo se enmascare en una mezcla formada mayoritariamente por compuestos poco activos, por lo que la medida de la actividad en promedio medida será baja y se descarte.

El método de mezclar y separar resulta útil en la preparación de bibliotecas de moléculas secuenciables (por ejemplo: péptidos de aminoácidos naturales), ya que las estructuras activas pueden finalmente identificarse aun teniendo muy poca cantidad de muestra. Para la síntesis de otras estructuras es necesaria la introducción de métodos de codificación, los cuales permitirán

la identificación del compuesto activo aún en pequeñas cantidades. Para ello pueden incluirse etiquetas en el mismo grano de la resina donde se realiza la síntesis de la molécula objetivo. En estos casos el proceso involucra la construcción simultánea de dos compuestos, por lo que la resina debe contener un grupo ligante con más de un sitio reactivo. Uno de ellos es la estructura objetivo y el otro es una etiqueta (usualmente un péptido u oligonucleótido) que permitirá deducir la composición de la molécula objetivo. En el esquema 3.12 se muestra un ejemplo que ilustra la síntesis de un compuesto que tiene 3 bloques constructores, por lo que al final de la reacción la etiqueta será un tripéptido cuya identidad definirá a los sustituyentes R, R' y R'' de la molécula objetivo. Luego de la síntesis la estructura no peptídica se puede separar selectivamente de la resina con los reactivos adecuados, dejando unida la etiqueta peptídica a la resina, la cual puede ser secuenciada para la identificación del producto liberado. Del mismo modo puede usarse un oligonucleótido como etiqueta, el cual puede amplificarse mediante alguna técnica adecuada.

Esquema 3.12: Síntesis de bibliotecas con identificación por etiquetas moleculares.

4.3 Bibliotecas de compuestos en paralelo

Como se mencionó anteriormente, las técnicas de síntesis en paralelo implican reacciones simultáneas en recipientes de reacción separados. Como ejemplo se mencionará el sistema Multipin, un método manual desarrollado por Geysen en 1984 (Figura 3.8). Consiste en un dispositivo

formado por varillas de polietileno insertadas sobre una placa, en cuyo extremo libre se encuentran inmovilizadas las resinas de poliestireno-divinilbenceno funcionalizadas con un grupo que actúa de espaciador y que posee una función amino libre capaz de generar reacciones de acoplamiento. Los reactivos se colocan en reservorios cuyo espaciado coincide con el de las varillas. En el caso de la síntesis de péptidos, por ejemplo, en los reservorios se ubican aminoácidos protegidos y se sumergen las varillas de modo de generar el primer acoplamiento a la resina. Luego se lavan las varillas y se realiza la remoción de los grupos protectores, para posteriormente repetir el proceso colocando otros aminoácidos protegidos en los reservorios. La etapa final consiste en la liberación del péptido, la cual puede realizarse en los mismos reservorios que sirven como recipientes adecuados para los ensayos biológicos.

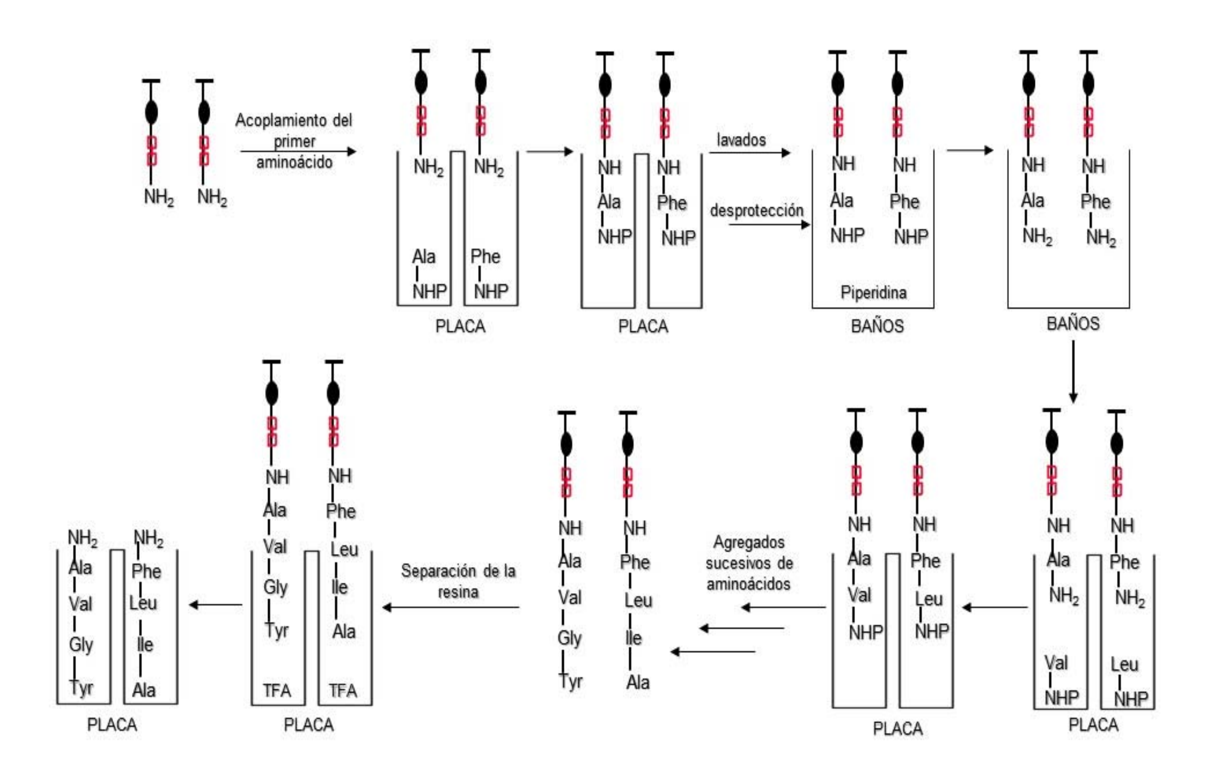


Figura 3.8: Representación de la síntesis en paralelo por el sistema Multipin. El grupo espaciador se simboliza en rojo. Las siglas NHP representan un grupo amino protegido como extremo del aminoácido.

Una alternativa a este dispositivo es el Diversomero®, desarrollado por Parke-Davis. A diferencia del sistema Multipin, donde la resina se une a una varilla por su parte exterior, en este caso la resina se encuentra en el interior de varillas huecas pero permeables al medio de reacción mediante un cierre de vidrio poroso. Esto permite adaptar una cámara para la refrigeración de las varillas en su parte superior, posibilitando reacciones a reflujo. Además, es posible sellarlas, permitiendo generar atmósfera inerte en los medios de reacción.

4.4 Diseño de quimiotecas.

Los primeros desarrollos en química combinatoria se han centrado en la síntesis de péptidos y nucleótidos, ya que son estructuras que pueden ser preparadas en fase sólida en forma automatizada mediante la repetición de unos pocos pasos de síntesis. Posteriormente se avanzó hacia la generación de bibliotecas de compuestos más pequeños, los cuales posiblemente resultan mejores candidatos a fármacos que los péptidos debido a la poca absorción oral y estabilidad de estos últimos. La construcción de estas nuevas bibliotecas implico un mayor desarrollo de la síntesis orgánica en fase sólida, dado que en estos casos debe recurrirse a otro tipo de reacciones para la generación de los productos finales.

Como resultado de la síntesis de una quimioteca se obtienen moléculas que comparten un esqueleto común (andamio o *scaffold*, en inglés) y que poseen diversos sustituyentes (Figura 3.9). Un scaffold ideal es aquel que posee un bajo peso molecular, ya que los productos finales tendrán sustituyentes que aumenten su tamaño. Además, resulta más conveniente que los sustituyentes se ubiquen en forma radial alrededor del esqueleto y no estén restringirlos a una pequeña región de la molécula. Finalmente se espera que un esqueleto contenga grupos funcionales que permitan reacciones de variación de sustituyentes en forma independiente para cada sitio reactivo. Ejemplos de esqueletos son los benzodiacepinas, las hidantoínas, los bifenilos, etc.

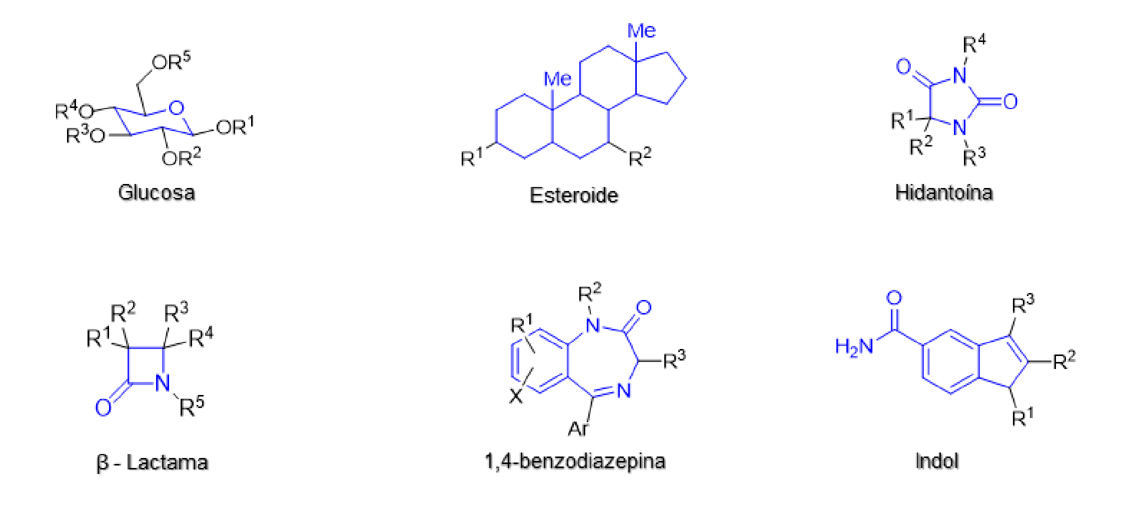


Figura 3.9: Estructura de scaffolds utilizados en la síntesis de quimiotecas.

5. Síntesis orgánica asistida por microondas

La técnica de síntesis asistida por microondas (MAOS) ha sido ampliamente reconocida como una importante herramienta alternativa para proporcionar energía térmica a una reacción, que presenta numerosas ventajas respecto del calentamiento térmico. Entre ellas se encuentran: el ahorro de tiempo y energía, la optimización de rendimientos, la obtención de productos en un

menor número de pasos y la generación de menos subproductos minimizando la contaminación. Además, en muchos casos se prescinde del uso de disolventes. Esta unión entre la ausencia de disolventes y las microondas, así como el tiempo requerido en las reacciones químicas alcanza un gran potencial en la síntesis orgánica y, sobre todo, en la síntesis de fármacos en forma automatizada y/o mediante técnicas de química combinatoria.

La radiación por microondas, como cualquier radiación electromagnética, está constituida por un campo eléctrico y un campo magnético (aunque sólo el campo eléctrico transfiere la energía que conduce al calentamiento de las sustancias). Están situadas en la región entre las ondas de radio y el infrarrojo con una longitud de onda comprendida entre 1cm y 1m que corresponde a frecuencias entre 30 GHz y 300 MHz. Debido a que los radares y otros sistemas de telecomunicaciones utilizan frecuencias ubicadas en la zona de microondas, para evitar interferencias existe una regulación a nivel nacional e internacional en la cual se utilizan bandas permitidas ISM (frecuencias industriales científicas y médicas), que para el uso de reactores microondas con fines químicos deben operar a una frecuencia de 2.45 GHz (longitud de onda de 12.25cm).

La energía generada por la radiación de microondas en la frecuencia de 2,45 GHz es relativamente baja (1,6 x 10^{-3} eV). Este nivel de energía es mucho más bajo que el nivel de energía de un enlace covalente o de hidrógeno (0.04 a 0.44 eV) lo que indica que la energía de microondas a esta frecuencia no puede romper estos enlaces. Por lo tanto, Las microondas no pueden inducir reacciones químicas por absorción directa de energía electromagnética a diferencia de la radiación ultravioleta y visible. Entonces *¿Cómo afecta la irradiación de microondas a una reacción química?* La respuesta está en el calentamiento eficiente (calentamiento dieléctrico) que generan las moléculas por transformar energía electromagnética en calor a través de dos mecanismos principales: polarización dipolar y conducción iónica.

5.1 Polarización dipolar y conducción iónica

Las moléculas con momento dipolar permanente o inducido, bajo la influencia de un campo eléctrico externo, tienden a alinear su momento dipolar con éste (Figura 3.9). Como el campo aplicado oscila, los dipolos tienden a alinearse al mismo. Así, se genera una diferencia de fase entre el dipolo y la orientación del campo. Esta diferencia de fase produce una pérdida de energía en forma de calor por fricciones, colisiones moleculares y por pérdidas dieléctricas, dando lugar al calentamiento dieléctrico por el mecanismo de rotación molecular. La cantidad de calor generada es directamente proporcional a la capacidad de los dipolos de alinearse con el campo aplicado. Es por esto que los gases no pueden calentarse mediante irradiación de microondas, debido a que las distancias entre las moléculas en rotación son demasiado grandes y no poseen, además, momentos dipolares adecuados para las pérdidas dieléctricas.

La segunda forma de transferencia de energía es la conducción iónica, que tiene lugar si hay iones libres o especies iónicas presentes en la sustancia. Según este mecanismo, el calor se genera a través de pérdidas por fricción, que resultan de la migración de iones disueltos cuando sobre ellos

actúa un campo electromagnético (Figura 3.9). Dichas pérdidas dependen del tamaño, carga, conductividad de los iones disueltos e interacción de estos últimos con el disolvente.

Los efectos de la conducción iónica son especialmente importantes cuando se trabaja con líquidos iónicos en microondas. En este caso, el mecanismo de conducción iónica es mucho más fuerte que el de rotación dipolar en cuanto a lo que a la generación de calor se refiere. Asimismo, el mecanismo de conducción iónica permite explicar el hecho de que algunos metales cuyo momento dipolar es nulo, se calientan efectivamente con microondas.

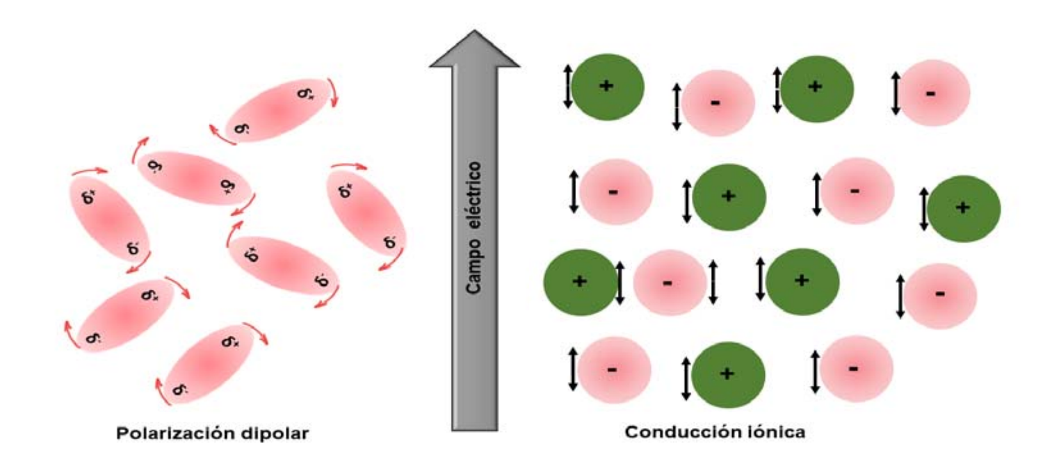


Figura 3.9: Ilustración esquemática de los mecanismos que generan calentamiento dieléctrico: Polarización dipolar (los dipolos se alinean con el campo eléctrico) y conducción iónica (los iones migran en dirección al campo eléctrico).

Los mecanismos de calentamiento mencionados anteriormente necesitan de que el acoplamiento entre las componentes del material irradiado y el campo eléctrico de las microondas sea efectivo. Ambos están influenciados por varios factores, entre ellos la longitud de onda de la radiación, las propiedades físicas de la sustancia (constante dieléctrica, polaridad, temperatura) y las características del ión (tamaño, concentración, carga y movilidad).

5.2 Propiedades dieléctricas

Para cuantificar la capacidad que tiene una sustancia de transformar la radiación de microondas en calor se utiliza el factor de disipación (tan) que indica la susceptibilidad a la radiación microondas. Este factor se define como el cociente entre la pérdida dieléctrica (”), que muestra la eficiencia con que la radiación electromagnética se transforma en calor, y la constante dieléctrica (’) que describe la polarizabilidad de la molécula en un campo eléctrico (Ecuación 5.1).

$$\tan \delta : \varepsilon'' / \varepsilon'$$

Ecuación 5.1: Expresión matemática del factor de disipación.

Para llevar a cabo una síntesis cuando las propiedades dieléctricas de la mezcla de reacción son demasiado pobres para permitir un calentamiento eficiente por radiación de microondas, la adición de pequeñas cantidades de aditivos (por ejemplo, sales o líquidos iónicos) que tienen grandes valores de tan pueden superar significativamente estos problemas y posibilitan un adecuado calentamiento de la mezcla. A su vez, es necesario determinar la necesidad del uso o no de disolvente. La posibilidad de no usar disolventes para las reacciones con microondas es una opción que en los últimos tiempos ha ido ganando adeptos siguiendo la filosofía de la *química verde*, permitiendo reacciones más respetuosas con el medio ambiente y reduciendo costos.

Para el caso en que sea indispensable el empleo de un disolvente, un factor crucial para su elección es su polaridad. Cuando los reactivos no absorben la radiación microondas, se ha de utilizar un disolvente polar (ya sea de polaridad intermedia o elevada). Cuanto mayor es la polaridad de un disolvente mayor es su capacidad para acoplarse con las microondas permitiendo un aumento más rápido de la temperatura interna. En la tabla 5.1 se puede observar una clasificación de disolventes orgánicos de uso habitual en función de su facilidad para acoplarse con las microondas.

Tabla 1. Solventes orgánicos de uso habitual clasificados según su eficiencia de calentamiento (tan δ), en microondas.

Elevado tan δ > 0.5		Medio tan δ 0.1-0.5		Bajo tan δ <0.1	
Solvente	**tan δ**	**Solvente**	**tan δ**	**Solvente**	**tan δ**
Etilenglicol	1.320	2-butanol	0.447	Cloroformo	0.091
Etanol	0.941	Diclorobenceno	0.280	Acetonitrilo	0.062
DMSO	0.825	Ácido acético	0.174	Acetato de etilo	0.059
2-propanol	0.799	DMF	0.161	Acetona	0.054
Metanol	0.659	Dicloroetano	0.127	Diclorometano	0.042
nitrobenceno	0.589	Agua	0.123	Tolueno	0.040
1-Butanol	0.571	Clorobenceno	0.101	Hexano	0.020

El empleo de disolventes apolares (ejemplo: hexano, benceno, tolueno), que no se acoplan de manera efectiva con las microondas, evita una elevación excesiva de la temperatura del medio de reacción. Esta característica es interesante para trabajar con sustancias sensibles a temperaturas elevadas, ya que estos disolventes disipan el calor que se produce cuando las microondas interaccionan con los reactivos polares. A continuación, se presentan algunos ejemplos que destacan las capacidades únicas y potenciales de la síntesis asistida por microondas en el campo de la química medicinal.

5.3 MAOS en el proceso de descubrimiento de fármacos

La síntesis asistida por microondas tiene influencia en varias etapas del descubrimiento de fármacos y no se limita solo a áreas relacionadas con la síntesis orgánica. Por ejemplo, la tecnología de microondas también se está utilizando en el descubrimiento de blancos moleculares. Los tiempos de reacción extremadamente cortos facilitados por MAOS permiten sintetizar compuestos de vida media corta que no se podrían obtener de otro modo utilizando métodos clásicos. Un ejemplo de esta aplicación es la preparación de radio fármacos que contienen isótopos con tiempo de vida media cortas (Figura 3.10; por ejemplo, ^{13}C, ^{3}H y ^{18}F). En estos ejemplos, esta técnica ha permitido con éxito reducir el tiempo de reacción, y aumentar el rendimiento del producto final (hasta duplicarlo en algunos casos).

Figura 3.10: Compuestos marcados preparados por síntesis orgánica asistida por microondas.

Andersen han aplicado está técnica en conjunto con catalizadores de metales de transición en la síntesis de Tripanavir, un fármaco utilizado en combinación con el Ritonavir en la terapia antiretroviral en pacientes infectados con VHI-1. Como se puede ver en el Esquema 3.10, las reacciones se llevaron a cabo usando los mismos pasos que en la síntesis tradicional, sin embargo, los tiempos de reacción para producir un buen rendimiento se redujeron desde horas a tan solo unos pocos minutos.

Esquema 3.10: Alquilación asimétrica catalizada por molibdeno bajo irradiación microondas en la síntesis de Tripanavir.

A los cortos tiempos de reacción se suma la alta tasa de selectividad de las reacciones que utilizan este medio alternativo de calentamiento, las que han abierto una puerta en la generación de numerosas bibliotecas de nuevos compuestos. Un ejemplo de ello es el que se presenta en el Esquema 3.11, donde se obtuvieron sub-bibliotecas a partir de scaffolds de dihidropirimidinas utilizando una gran diversidad de reacciones clásicas.

Esquema 3.11: Esqueleto de dihidropirimidina utilizado para crear 9-sub-bibiotecas a partir de las siguientes reacciones: a) aminocarbonilación catalizada por Pd (X=Br), b) N- arilación catalizada por Cu, c) N-acilacion con cloruros de ácido o anhídridos, d) reacción del tipo Liebeskind-Srogl catalizada por Pd (Z<C=>S), e) N-alquilación de Mitsunobu, f) cicloadición de azida-alquino catalizada por Cu (después de la bromación C6 y el desplazamiento con azida), g) alquilación de Mitsunobu (RX=OH), h) reacción de Liebeskind-Srogl catalizada por Pd(RX=EtS), i)Formación de enlaces amida(RX=OH).

Referencias

Furlán, R. et al (2012). Química Combinatoria: Metodologías relacionadas con la generación de la diversidad molecular. México: Fondo de cultura económica.ISBN 978-607-16-0670-9

Kappe, C. O. et al (2005) Microwaves in Organic and Medicinal Chemistry.5th ed. Wiley-VCH.

Kappe,O. (2018). My Twenty Years in Microwave Chemistry: From Kitchen Ovens to Microwaves that aren't Microwaves. *The chemical Record*, 18, 1–26.

Patrick, G. (2015). An Introduction to Drug Synthesis. Oxford University Press.

CAPÍTULO 4
Descriptores moleculares

Melisa E. Gantner

1. Descriptores moleculares

Un descriptor molecular es una variable numérica, discreta o continua, que refleja distintos aspectos de una estructura molecular. Los investigadores Todeschini y Consonni brindan la definición más aceptada hasta el momento: *"Los descriptores moleculares son el resultado final de un procedimiento lógico-matemático el cual transforma la información química codificada dentro de una representación simbólica de una molécula en un número útil o el resultado de un experimento estandarizado"*.

El número de descriptores moleculares propuestos hasta el momento en la literatura es realmente amplio. Básicamente, estos pueden ser de carácter empírico o teórico. Los descriptores empíricos derivan de mediciones experimentales, mientras que los teóricos derivan de una representación simbólica de la molécula. Estos últimos son particularmente útiles debido a que no requieren ninguna determinación experimental para caracterizar las moléculas y, como veremos más adelante, algunos descriptores experimentales también se pueden obtener mediante aproximaciones teóricas. Adicionalmente, un descriptor molecular puede describir a la molécula como un todo (descriptores globales) o solo representar un fragmento presente en ella.

Cada representación molecular constituye una forma diferente de mirar, interpretar y extraer información de la estructura de una molécula. Esencialmente, los descriptores moleculares pueden clasificarse según la dimensionalidad de la representación molecular de la cual derivan en:

- descriptores independientes de la conformación espacial y/o de la orientación en el espacio de la molécula y,
- descriptores dependientes de la conformación espacial y/o de la orientación en el espacio de la molécula.

1.1 Descriptores independientes de la conformación

En este caso, el valor que asume el descriptor no depende del confórmero de la estructura a partir de la cual se calculan, ni de la orientación del mismo en un marco de referencia determinado. Se los conoce como descriptores de baja dimensionalidad (0-2D). La independencia de la

conformación espacial genera una de las principales ventajas que conlleva el uso de este tipo de descriptores que es su bajo costo computacional, dado que no requieren de un análisis conformacional previo de las estructuras químicas (optimizando notablemente los tiempos de cálculo empleados). Este grupo engloba varios subgrupos de descriptores moleculares tales como:

- Descriptores 0D: descriptores constitucionales derivados de la fórmula molecular. Son independientes de cualquier conocimiento sobre la estructura molecular. Describen solamente la constitución de la molécula, pero no dicen nada sobre la conformación ni tipo de conectividad presente. Los más simples son el peso molecular, número y tipo de átomos, suma de propiedades atómicas, el número de enlaces, entre otros.

- Descriptores 1D: describen fragmentos de las moléculas constituidos por el agrupamiento de sus átomos constituyentes (lista de fragmentos estructurales de la molécula) y no requieren del conocimiento completo de la estructura molecular. De esta manera, son variables que señalan la presencia/ausencia o frecuencia de un tipo de átomo, grupo funcional o subestructura/fragmento determinado de la molécula y asumen valores binarios (en el caso que indiquen ausencia o presencia) o en su defecto valores no binarios pero discretos (cuando indican la frecuencia de aparición).

Algunos autores incluyen en este subgrupo a descriptores globales como las propiedades fisicoquímicas simples que surgen de mediciones experimentales estandarizadas: solubilidad, valor de pKa, coeficiente de partición octanol/agua (logP), índice de refracción, entalpía de vaporización, punto de ebullición, volumen molar, etc.

El coeficiente de partición octanol/agua (logP) ha sido la medida de lipofilicidad elegida en el desarrollo de moléculas activas, para las cuales el transporte a través de las membranas biológicas es generalmente crítico. Además, es el parámetro fisicoquímico más popular para definir la lipofilicidad de compuestos en los estudios estructura-actividad. Formalmente, el logP se define como el logaritmo decimal de la relación de las concentraciones en equilibrio de una sustancia disuelta en un sistema bifásico consistente en dos disolventes considerablemente inmiscibles, usualmente n-octanol y agua (Ecuación 4.1).

$$logP = log\left(C_{n-octanol}/C_{agua}\right)$$

Ecuación 4.1: Expresión matemática general que define al coeficiente de partición.

Adicionalmente el valor del logP se puede determinar experimentalmente mediante cromatografía en capa fina de fase reversa o HPLC de fase reversa. Esta propiedad ha sido medida experimentalmente para un gran número de compuestos orgánicos, pero, en términos prácticos, resulta imposible medir el logP para todos los compuestos químicos existentes. Es por esto que predecir este valor (incluso para compuestos aún no sintetizados) es de suma utilidad. En este sentido se han propuesto varios modelos teóricos. El primero de ellos se atribuye a Hansch y colaboradores debido a la introducción de la constante   de Hansch, la cual otorga un valor

numérico a la influencia de los sustituyentes de un compuesto sobre el valor total del logP. Según esta aproximación, el logP puede ser calculado considerando la contribución a la hidrofobicidad de cada sustituyente a la molécula sin sustituir (o *parent*), y cada contribución se define como el parámetro de lipofilia del sustituyente (Ecuación 4.2).

$$logP = logPparent + \sum \pi$$

Ecuación 4.2: Estimación del logP mediante el método de Hansch y colaboradores.

Donde para el sustituyente *x* se calcula de acuerdo con la Ecuación 4.3:

$$\pi_x = logPHx - logPH$$

Ecuación 4.3: Cálculo de la contribución a la lipofilia de un sustituyente x.

Siendo H la molécula sin el sustituyente *x*. Por ejemplo, el valor para el sustituyente metilo aromático se puede calcular de acuerdo a la Ecuación 4.4:

$$\pi_{metilo} = logPmetilbenceno - logPbenceno$$

Ecuación 4.4: Ejemplo del cálculo del parámetro π.

Otro método de cálculo teórico del logP se conoce como método de Rekker (Ecuación 4.5), según el cual el cálculo está basado en la descomposición de la molécula en pequeños fragmentos que poseen un valor teórico de lipofilicidad (*f*).

$$logP = \sum a_i f_i + \sum b_j f_j$$

Ecuación 4.5: Estimación del logP mediante el método de Rekker.

En este caso el cálculo consiste en la resolución de dos sumatorias, una contiene la suma de la contribución a la lipofilicidad de los fragmentos y otra la suma de términos de corrección. En

la Ecuación 4.5 el valor *a* corresponde al número de veces que está presente el fragmento *f* y *b* al número de veces que se corrige el factor *f*. La suma de estos fragmentos más los términos de corrección por efectos estéricos, electrónicos y de puentes de hidrógeno, da como resultado el logP de la molécula. El valor de *f* se calcula mediante métodos estadísticos basados en el análisis de fragmentos procedentes de moléculas con logP conocido.

Existen otras aproximaciones teóricas para el cálculo del logP, entre ellas el método de Leo, el método de Moriguchi, etc. Todas estas aproximaciones funcionan en general muy bien para compuestos simples, pero suelen presentar limitaciones para estructuras complejas por lo que no debe olvidarse que este tipo de cálculos son aproximaciones que permiten tener una estimación del logP.

Muchos otros descriptores fisicoquímicos también se pueden calcular mediante aproximaciones teóricas como por ejemplo descriptores de efectos electrónicos como la constante electrónica de Hammet ; descriptores de efectos estéricos como el factor estérico de Taft (Es), el parámetro de Verloop y la refractividad molar; entre otros.

Descriptores 2D: se basan en la representación bidimensional o topológica de la molécula y por este motivo se los conoce como descriptores topológicos.

La topología puede definirse como aquella parte del álgebra que estudia las posiciones e interconexiones de los elementos dentro de un conjunto. Aplicada a las moléculas, ha dado origen a una nueva disciplina llamada “topología molecular” o Teoría de Grafos Química, que analiza las posiciones relativas y la conectividad de los átomos dentro de las moléculas. Desde este punto de vista no se aborda el estudio de aspectos tales como la estructura tridimensional del compuesto, los tipos de enlaces, ángulos entre ellos y, a veces, ni siquiera la naturaleza de los átomos enlazados. La cuestión fundamental recae en qué átomos están ligados entre sí y cuáles caminos conducen de un átomo a otro dentro de la molécula. De esta manera los descriptores topológicos proporcionan información de constitución y conectividad. El método topológico emplea los llamados “índices topológicos” para caracterizar estructuralmente a un compuesto. Para definir dichos índices, el primer paso es representar a los átomos por puntos (vértices) y a los enlaces por segmentos (ejes o aristas), eliminándose ocasionalmente los átomos de hidrógeno. De esta manera, se obtiene el “grafo” de la molécula. Posteriormente, se numeran aleatoriamente los distintos átomos (o vértices) y se construye alguna “matriz topológica” (por ejemplo, matriz distancia, matriz adyacencia, matriz de distancias topológicas máximas, etc). A partir de esta matriz y mediante operaciones algebraicas más o menos sencillas se obtienen los “índices topológicos” más importantes, que pueden caracterizar aspectos muy diversos de la estructura molecular tales como la forma, el grado de ramificación o ciclación de una estructura química o la distribución de cargas (Figura 4.1).

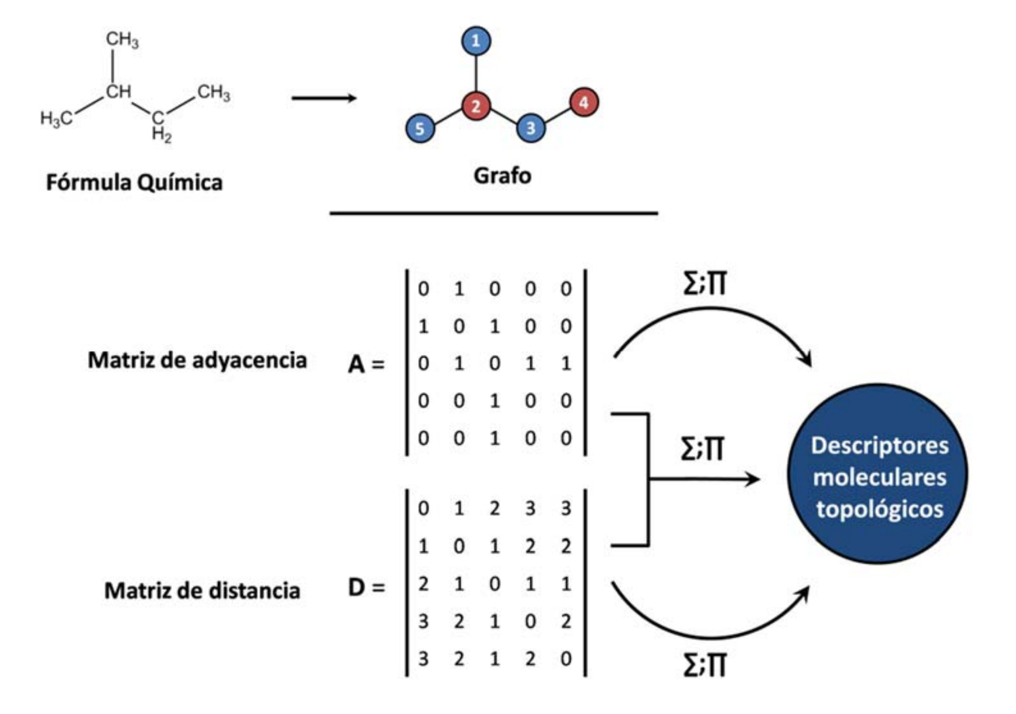


Figura 4.1: Esquema simplificado del cálculo de descriptores topológicos. La estructura química (isopentano en este caso) es inicialmente representada como un grafo a partir del cual se pueden construir matrices, como la matriz de adyacencia y la de distancia (dos de las matrices de menor complejidad en lo que respecta al cálculo de índices topológicos), de las que finalmente surge un valor denominado descriptor molecular topológico.

Ejemplos representativos de este tipo de descriptores incluyen los índices de Conectividad de Wiener y Rándic. El índice de Wiener (W) se define de acuerdo con la Ecuación 4.6.

$$W = 1/2 \sum_{ij} d_{ij}$$

Ecuación 4.6: Índice de conectividad de Wiener

En la Ecuación 4.6, d_{ij} representa la distancia topológica entre los vértices v_i y v_j si se considera el camino de longitud más corta. La longitud u orden del camino es el número de aristas que lo componen.

Por su parte, el índice de Rándic (*X*) se define de acuerdo a la Ecuación 4.7:

$$X = \sum_{todos\ los\ ejes} \left(deg_i \cdot deg_j\right)^{-1/2}$$

Ecuación 4.7: Índice de Rándic.

En dicha ecuación, para cada arista (v_i, v_j), el deg_i es el grado de degeneración del vértice v_i y representa el número de vértices adyacentes al mismo.

Una de las desventajas de los descriptores topológicos es su naturaleza compleja y abstracta, lo que los hace difíciles de interpretar desde el punto de vista fisicoquímico y en relación con los aspectos biológicos. Por otra parte, es imposible para un descriptor independiente de la conformación diferenciar isómeros de geometría, aunque sabemos, sin embargo, que las propiedades fisicoquímicas y biológicas de éstos pueden ser muy distintas. Otra desventaja es que poseen alto grado de degeneración, lo que significa que muchas veces poseen un mismo valor numérico para estructuras diferentes. Esto se debe a que los índices topológicos derivan de representaciones algebraicas del grafo (matrices) que sólo tienen en cuenta la conectividad entre los átomos y no consideran la naturaleza de los mismos o de los enlaces covalentes que los unen. Así, por ejemplo, ninguno de los índices previamente mencionados podría diferenciar a la anilina del fenol, ya que ambos compuestos poseen grafos isomorfos. Para hacerlo se utiliza la estrategia de ponderar de distinta manera los vértices y aristas del grafo según las propiedades químicas de los átomos y enlaces involucrados, dando lugar a un nuevo conjunto de matrices ponderadas que permiten derivar un nuevo grupo de descriptores topológicos cuyo tratamiento escapa al contenido de este libro.

1.2 Descriptores dependientes de la conformación

En este tipo de descriptores el valor que asume el descriptor sí depende de la conformación y/o en algunos casos, de la orientación de la molécula en relación a un marco de referencia. Este tipo de descriptores suele denominarse de alta dimensionalidad y, como veremos a continuación, si bien su cálculo es más complejo y costoso computacionalmente con respecto a los de baja dimensionalidad, permiten considerar tautómeros, estereoisómeros y los distintos estados de protonación de las moléculas en estudio. En esta categoría agrupamos a los descriptores 3D y 4D.

Descriptores 3D: derivan de la representación tridimensional de la molécula y se basan no solo en la naturaleza y conectividad de los átomos sino también en la conformación y configuración espacial. Ejemplos de descriptores 3D son los descriptores de volúmen (volúmen de van der Waals, volúmen geométrico), 3D-MoRSE (*3D Molecule Representation of Structure based on Electron diffraction*), RDF (*Radial Distribution Function*), WHIM (*Weighted Holistic Invariant Molecular descriptors*), GETAWAY (*GEometry, Topology, and Atom-Weights AssemblY*), entre muchos otros.

Debido a la dependencia de la conformación, el cálculo de los descriptores 3D se realiza utilizando, al menos, una conformación energética probable que generalmente corresponde a un mínimo local de la energía potencial de la molécula. Es importante destacar aquí que una conformación de mínima energía no debe confundirse con la conformación bioactiva, que es aquella conformación que adopta la molécula al interaccionar con un receptor biológico.

La minimización de la energía de un sistema molecular se lleva a cabo por cálculos iterativos sucesivos en los que se somete a la conformación inicial del sistema a un proceso de optimización geométrica parcial o completa, que consiste en pequeños cambios en los parámetros (todos o algunos) que definen la geometría del sistema, hasta que la estructura alcanza un mínimo energético local. Para este proceso se utilizan algoritmos de minimización que buscan matemáticamente los mínimos de la función de energía potencial del sistema, la cual es calculada por un método adecuado, ya sea un método basado en campos de fuerza de la mecánica molecular (por ejemplo, MM+) o métodos derivados de la mecánica cuántica (*ab initio* o semiempíricos), los cuales se verán en detalle más adelante en este capítulo.

Existen numerosos algoritmos de minimización, algunos de ellos hacen uso de la derivada primera de la función de energía potencial (por ejemplo, los gradientes conjugados), o de la derivada segunda (por ejemplo, Newton-Raphson), para guiar la minimización hasta alcanzar el mínimo local más próximo. Ninguno de estos métodos puede garantizar que se ha encontrado la conformación de energía más baja posible, es decir, el mínimo global de energía potencial. Encontrar tal conformación implica realizar lo que se conoce como búsqueda conformacional, es decir, explorar completamente el espacio conformacional de una molécula y encontrar el mínimo global de energía entre un enorme número de mínimos locales presentes en la hipersuperficie de energía potencial característica de esa molécula (Figura 4.2).

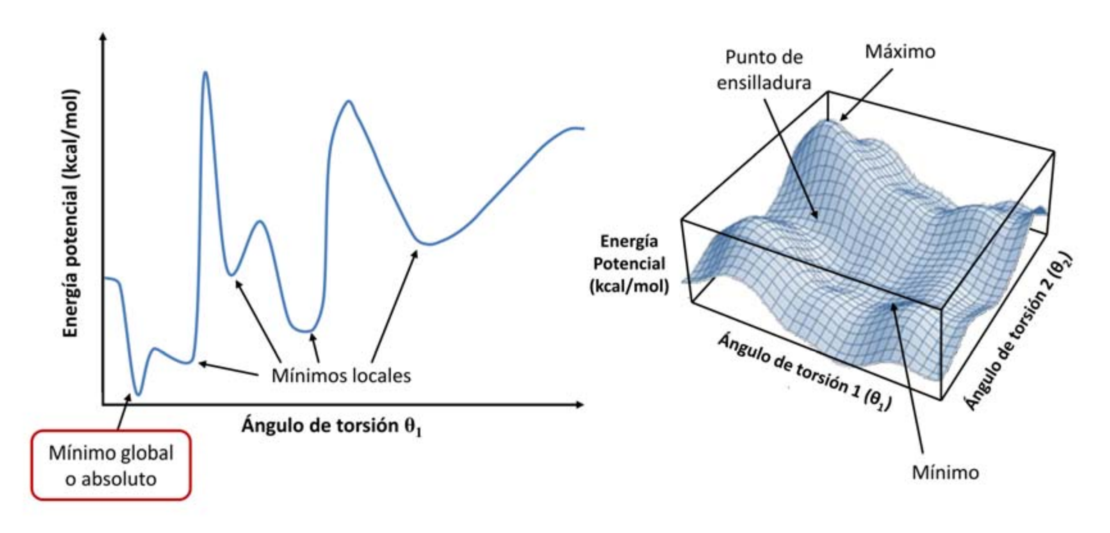


Figura 4.2: Gráficas de energía potencial de un sistema que posee uno o dos ángulos de torsión.

Los ángulos de torsión (también llamados ángulos diedros) de una molécula están definidos por 4 átomos unidos entre sí por 3 enlaces. El número de ángulos de torsión presentes es una molécula, determina el grado de libertad rotacional de la misma. En la imagen de la izquierda de la Figura 4.2 se representa en forma bidimensional a la variación de la energía potencial de una molécula en función de un único ángulo de torsión (θ_1). En ella se pueden observar varios máximos y mínimos locales y también el mínimo global o absoluto. A la derecha, la variación de la

energía potencial se representa en función de dos ángulos de torsión (θ_1 y θ_2), resultando un gráfico en tres dimensiones donde se observa una superficie de energía potencial, con sus mínimos, máximos y puntos de ensilladura. Así, una hipersuperficie de energía potencial es una gráfica de n+1 dimensiones, siendo n el número de grados de libertad del sistema, cuya forma ya no puede conocerse para un sistema de 3 o más ángulos de torsión.

Nos encontramos entonces frente a un problema con múltiples variables donde la respuesta que se obtiene depende en gran medida del punto de partida. El punto de partida lo puede constituir un confórmero cualquiera, que se optimiza mediante alguna técnica de minimización de la energía, y se compara con otros confórmeros encontrados previamente para chequear que no sea un duplicado. Si el nuevo confórmero así generado no había sido descubierto aún, se añade a una lista acumulativa de confórmeros distinguibles entre sí. Este ciclo continúa utilizando una nueva geometría de partida. Llegará un momento en que se habrán utilizado todas las geometrías iniciales posibles o en que no se encuentren nuevos mínimos, con lo que se dará por finalizada la búsqueda. Aquel confórmero de la lista con el valor de energía potencial más baja será considerado el confórmero de menor energía correspondiente al mínimo global de energía de esa molécula.

Como la etapa correspondiente a la minimización simplemente refina las estructuras de partida hasta alcanzar el mínimo local más próximo, quien más controla la eficacia de la búsqueda para alcanzar la convergencia son los métodos que generan los confórmeros iniciales. Éstos se pueden dividir en cuatro categorías: (1) búsquedas deterministas que cubren todas las áreas del espacio conformacional de forma sistemática (*grid search* o método sistemático), (2) métodos estocásticos (Monte Carlo) que utilizan un elemento aleatorio para la exploración del espacio, (3) algoritmos evolutivos y, (4) dinámica molecular.

1-Los métodos deterministas incluyen cualquier método por el cual la generación y evaluación de una conformación de una molécula está determinada por la conformación anterior. De esta manera, en una búsqueda sistemática se van variando cada uno de los ángulos de torsión de forma sistemática y se va minimizando la energía de cada uno de los confórmeros así generados. Por ejemplo, para una molécula con 3 ángulos diedros, y un incremento del valor de cada ángulo de 60°, se generarán 6^3 = 216 configuraciones que deben ser optimizadas. Si el número de torsiones considerados es 6, esta cifra aumenta a 46656 (6^6), y para 9 torsiones, el número de configuraciones generadas es superior a diez millones (6^9). Estos números sirven para poner de manifiesto que, al ser un procedimiento combinatorio, la demanda computacional restringe su uso a problemas con un número limitado de grados de libertad torsional.

2-Para sistemas que contienen muchos grados de libertad donde los métodos determinísticos son demasiado costosos (generalmente porque el espacio de búsqueda es demasiado grande para una búsqueda sistemática completa), suelen ser útiles los métodos estocásticos. Estos métodos están basados en la técnica de búsqueda de Monte Carlo, a partir de la cual se genera un conjunto de estructuras tridimensionales arbitrarias no relacionadas entre sí mediante la combinación de un muestreo al azar (de ahí que el nombre del método haga referencia al célebre casino) del movimiento total de la molécula y el empleo del factor de Boltzmann (e^{-E_i/k_BT}), siendo

E la energía potencial del sistema de la conformación i, k_B la constante de Boltzmann y T la temperatura. Así, dada una configuración de partida elegida arbitrariamente, se genera una nueva por desplazamiento al azar en el espacio cartesiano de uno o más átomos (o por variaciones en ángulos torsionales seleccionados), de forma tal que al cabo de un gran número de desplazamientos sucesivos se haya muestreado de forma uniforme el espacio disponible a todos los átomos. Cada nueva configuración se acepta o rechaza en base a un criterio energético (criterio de Metrópolis) que implica un cambio en la energía potencial del sistema (ΔE) con respecto a la configuración anterior: se acepta si $\Delta E \leq 0$, o si para $\Delta E > 0$, $e^{-\Delta \mathrm{E}/\mathrm{k_B T}} > R$, donde R es un número aleatorio tomado de una distribución uniforme sobre el intervalo $(0,1)$. En caso de ser rechazada la configuración actual, se vuelve a la configuración previa, que se utiliza de nuevo como punto de partida para otro desplazamiento al azar. Dado que en ciertos casos permite aceptar configuraciones con $\Delta E > 0$, este método tiene menos tendencia a quedar atrapado en mínimos locales.

3-Otros métodos estocásticos son los algoritmos evolutivos que copian las leyes de la evolución natural, donde una población evoluciona hacia el mejor candidato (confórmero de menor energía). Un ejemplo clásico de este tipo de algoritmo son los algoritmos genéticos, los cuales se analizarán en detalle en el Capítulo 6.

4-Finalmente, mediante la dinámica molecular (metodología que también se analizará en el Capítulo 6) puede realizarse la simulación del calentamiento del sistema, impartiéndole de ese modo energía para que barra todo el espacio conformacional. Luego el sistema se "enfría" a distintos tiempos de forma tal de caer en distintas zonas de la hipersuperficie de energía potencial, asociadas a distintas conformaciones estables (conformaciones de partida) que posteriormente se minimizan al mínimo local más próximo.

En la Figura 4.3 se ejemplifica este método. A la izquierda se grafican las etapas de una simulación por dinámica molecular. A partir de una temperatura inicial (T° de inicio), se calienta el sistema en un intervalo de tiempo establecido (tiempo de calentamiento) hasta la temperatura de simulación, la cual se mantiene durante un período de tiempo definido (tiempo de simulación), y finalmente se lleva el sistema a la temperatura final (T° final) en un tiempo también determinado (tiempo de enfriamiento). En el gráfico de la derecha, se representa esquemáticamente el fundamento de aplicar dinámica molecular para la generación de las conformaciones de partida. Durante este proceso, mediante el calentamiento se le entrega energía cinética al sistema de modo de barrer todo el espacio conformacional (flecha roja). Al producirse el enfriamiento a los distintos tiempos de simulación (flechas negras), el sistema "cae" en distintas regiones de la curva de energía potencial, asociadas a diferentes conformaciones estables, las cuales posteriormente se minimizan mediante un proceso de optimización de la energía hasta el mínimo local más cercano.

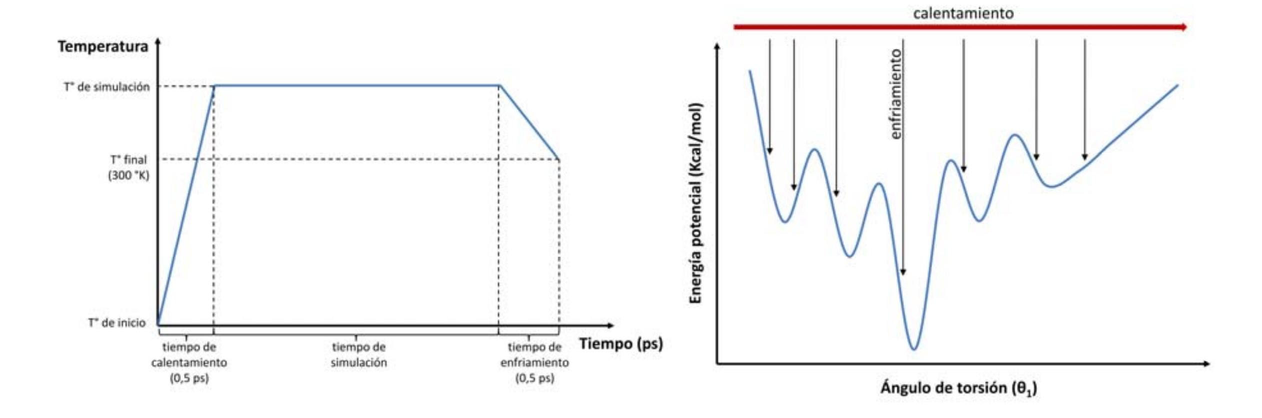


Figura 4.3: Representación del empleo de dinámica molecular como método de búsqueda de conformaciones.

Una desventaja de los métodos estocásticos es que son esencialmente heurísticos; dado que contienen un componente al azar no se puede controlar el proceso de forma de garantizar que puedan encontrar una solución óptima o simplemente una solución. Sin embargo, para muchos problemas estas técnicas han sido la única herramienta válida. Por otro lado, todavía no existe un fundamento teórico para decidir el criterio de convergencia de una búsqueda conformacional estocástica, en otras palabras, determinar cuándo el muestreo es suficiente, por lo que esta decisión es hecha empíricamente y es siempre dependiente del sistema en estudio y los objetivos de la búsqueda conformacional.

La finalidad de una búsqueda conformacional es encontrar las distintas conformaciones posibles y ordenarlas de acuerdo con sus energías, para seleccionar aquellas con las energías más bajas, que es de esperar que sean las que contribuyen a las propiedades del estado fundamental de la molécula. Sin embargo, no debe olvidarse que si las energías de estos mismos confórmeros se evaluaran en presencia de un disolvente polar, en lugar de *in vacuo*, es posible que un número significativo de interacciones intramoleculares serán reemplazadas por interacciones más favorables con moléculas del disolvente.

Descriptores 4D: examinan el espacio conformacional de los objetos moleculares, pero la información estructural se considera para un conjunto de confórmeros (condicionalmente, la cuarta dimensión), en lugar de una conformación fija como en el caso de los 3D. Ejemplos de este tipo de descriptores son aquellos derivados de los métodos GRID o CoMFA (*Comparative Molecular Field Analysis*).

En la Figura 4.4 se muestra un resumen esquemático de los descriptores moleculares y su clasificación según la dimensionalidad de la representación molecular de la cual derivan.

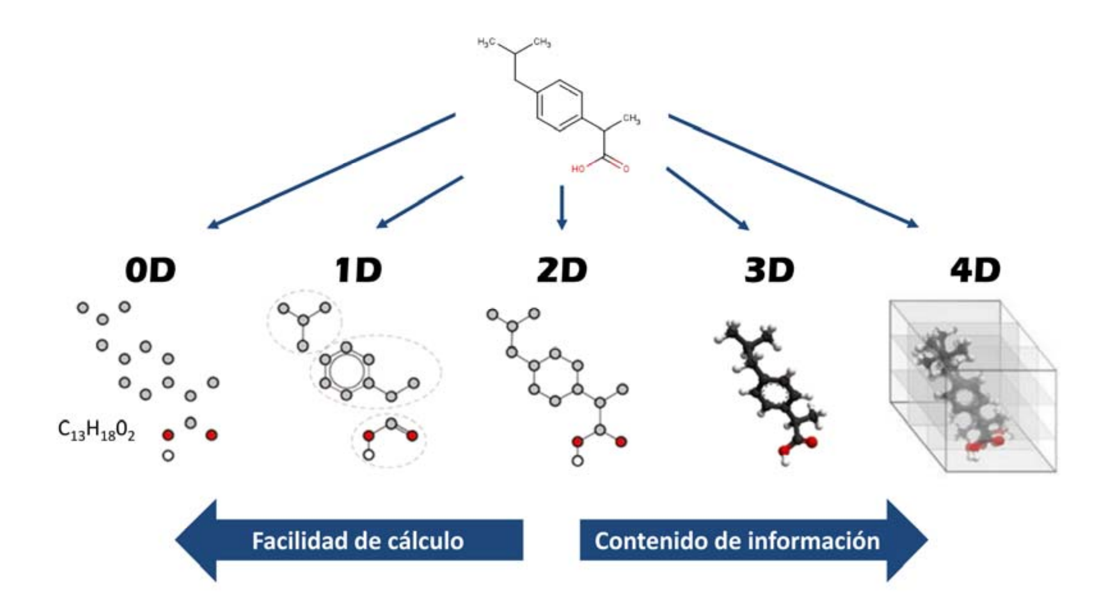


Figura 4.4: Clasificación de los descriptores moleculares según la dimensionalidad de la representación molecular de la cual derivan.

A modo de resumen puede decirse que los descriptores 0D derivan de la fórmula molecular y no requieren conocimiento de la estructura molecular; los 1D derivan de listas de fragmentos estructurales de la molécula y no requieren del conocimiento completo de la estructura molecular; los 2D derivan de la representación bidimensional de la estructura molecular conocida como grafo; los 3D derivan de la representación tridimensional de la molécula considerando para su cálculo un confórmero de menor energía; mientras que los 4D también derivan de una representación 3D de la molécula pero utilizan para su cálculo un conjunto de confórmeros en relación a un marco de referencia determinado. A medida que aumenta el contenido de información de un descriptor molecular, aumenta el costo computacional para su cálculo.

2. Cálculos de energía mediante química computacional

El modelado molecular asistido por computadora involucra el uso de programas o algoritmos para el cálculo de propiedades de las moléculas. Como hemos detallado previamente, es posible calcular diversos descriptores moleculares de origen teórico como por ejemplo la energía relacionada con una determinada conformación de una molécula. También hemos mencionado previamente que, en el caso de cálculos de descriptores dependientes de la estructura tridimensional de los compuestos, resulta conveniente modificar la conformación para llegar a un mínimo energético y calcular las propiedades en ese contexto. Los métodos computacionales que se utilizan para calcular la energía asociada a una conformación pueden dividirse en dos categorías: los métodos de origen clásico y los métodos de origen cuántico.

Los métodos de origen clásico se basan en la mecánica molecular, cuyas ecuaciones siguen las leyes de la física clásica y consideran a las moléculas como una colección de esferas (o

núcleos, representando a los átomos) unidas a través de resortes (representando a los enlaces). En esta aproximación los electrones no son considerados en forma explícita y la energía es calculada en función de las posiciones de los núcleos como una suma de términos. Usualmente la energía se expresa matemáticamente en lo que se denomina un campo de fuerza, en el cual se desglosan las contribuciones directas e indirectas.

Las contribuciones directas (o también denominadas enlazantes) tienen que ver con la energía asociada al estiramiento de enlaces, la oscilación angular y la rotación de los átomos a través de los ángulos de torsión. La Figura 4.5 muestra una representación de este tipo de interacciones. La interacción entre dos átomos que forman un enlace covalente puede aproximarse a un potencial armónico dependiente de la longitud de enlace r, la cual se desvía de una longitud de equilibrio r_0. Similarmente, puede ajustarse a un potencial de este tipo a la oscilación angular formada por tres átomos consecutivos que forman un ángulo θ que oscila alrededor de θ_0. También puede modelarse la energía asociada a la rotación de enlaces mediante la variación de ángulos diedros, la cual puede generar efectos estéricos entre átomos separados por 3 enlaces.

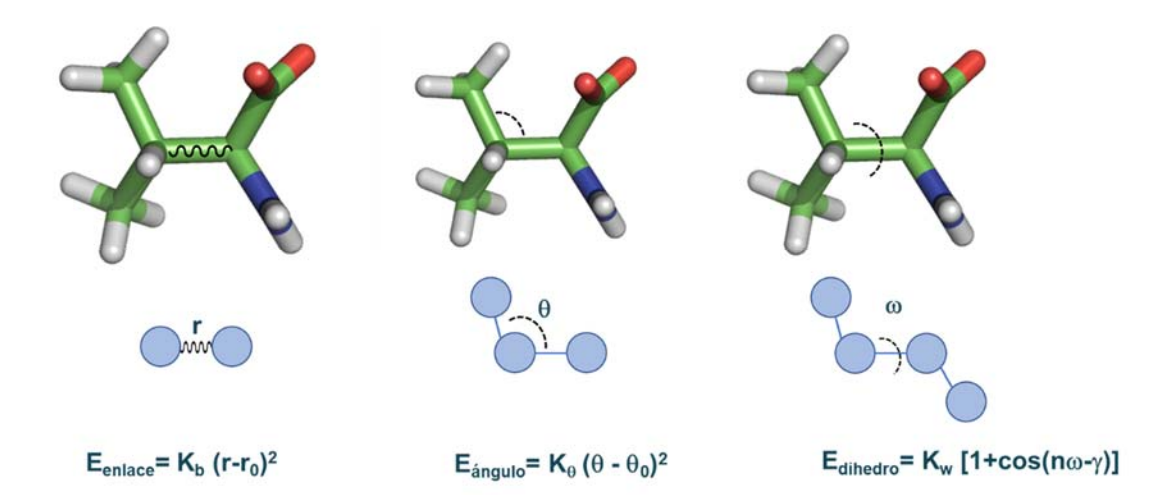


Figura 4.5: Representación de las interacciones enlazantes más importantes que contribuyen a un campo de fuerza de mecánica molecular. Se incluyen las expresiones matemáticas que pueden utilizarse para cuantificar cada contribución.

Las contribuciones indirectas (o no enlazantes) al campo de fuerza implican la interacción entre átomos no unidos entre sí. Entre estas últimas aparecen las interacciones electrostáticas que pueden representarse con funciones del tipo coulómbicas y las interacciones de Van der Waals, a las que puede asociárseles una forma potencial del tipo de Lennard-Jones (Figura 4.6).

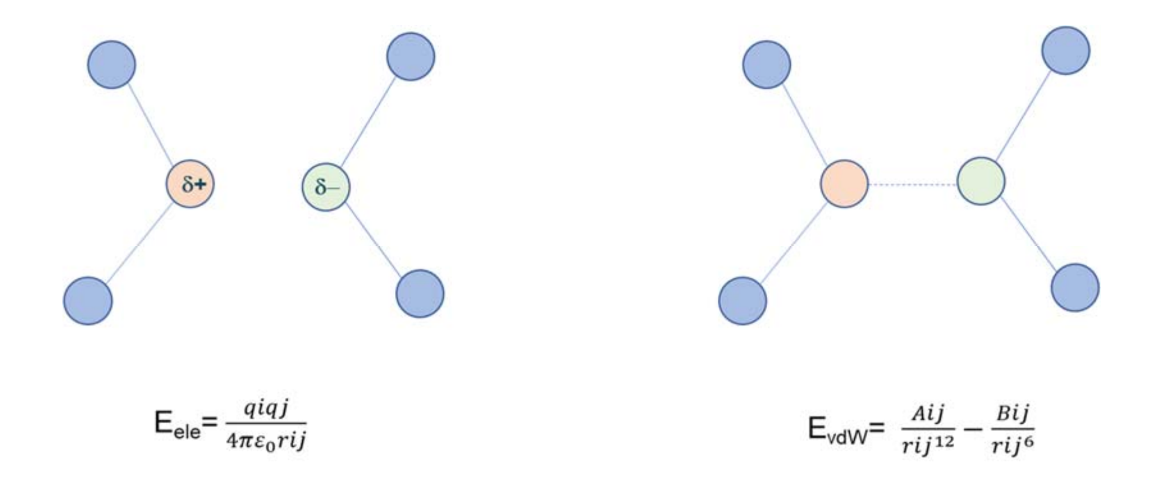


Figura 4.6: Representación de las interacciones no enlazantes más importantes que contribuyen a un campo de fuerza de mecánica molecular. Se incluyen las expresiones matemáticas que pueden utilizarse para cuantificar cada contribución.

Como se observa en las ecuaciones descriptas en las Figuras 4.5 y 4.6, los cálculos requieren información y parámetros por ejemplo r_0, θ_0, K, γ, A, B, etc. Estos valores dependerán del tipo de átomo que se considere en cada caso y se obtienen mediante métodos experimentales o cálculos mecano-cuánticos. Este conjunto de parámetros se encuentra tabulado en los programas y también forma parte del campo de fuerza. En otras palabras, los campos de fuerza contienen la ecuación matemática que representa a la energía, la definición de los tipos de átomos que se distinguen en el cálculo y sus parámetros asociados.

Los métodos que utilizan mecánica molecular son relativamente baratos en términos de costo computacional, por lo que pueden utilizarse para obtener valores de descriptores y son los métodos de elección cuando se requiere estudiar grandes moléculas como proteínas. En el Capítulo 6 describiremos su aplicación en el diseño y búsqueda de nuevos fármacos mediante técnicas de simulación. La gran limitación de este tipo de cálculos es que no es posible determinar propiedades electrónicas de las moléculas, debido a que los electrones no están incluidos explícitamente en esta aproximación.

Los métodos de origen cuántico utilizan la mecánica cuántica para el cálculo de las propiedades y energía de una molécula, considerando las interacciones núcleo-electrón del sistema. Con esta metodología se representan explícitamente los electrones en los cálculos, por lo que es posible derivar propiedades que dependen de la distribución electrónica y, por ejemplo, investigar reacciones químicas en las cuales se producen roturas o formación de enlaces.

El punto de partida es la ecuación de Schrödinger, la cual en situaciones en que la energía potencial es independiente del tiempo tiene la forma de la Ecuación 4.8 para una partícula de masa m que se mueve a través del espacio bajo la influencia de un potencial V.

$$\left\{\frac{-h^2}{8m\pi^2}\nabla^2 + V\right\}\Psi_{(r)} = E\,\Psi_{(r)}$$

Ecuación 4.8: Forma funcional de la ecuación de Schödinger independiente del tiempo para una partícula de masa m que se mueve en el espacio con un potencial V.

En la Ecuación 4.8, E representa a la energía de una partícula, por ejemplo, un electrón y Ψ es la función de estado o función de onda, la cual contiene toda la información significativa del sistema. La función de onda no tiene sentido físico directo. Sin embargo, el valor de su cuadrado (Ψ^2) en un punto del espacio representa la densidad de probabilidad de encontrar al electrón en dicho lugar; y el producto Ψ^2 dv, la probabilidad de encontrarlo dentro del volumen infinitesimal dv.

El lado izquierdo de la ecuación puede simbolizarse con un operador, en particular con el operador Hamiltoniano: $\hat{H}$. Esto reduce la expresión a $\hat{H}\Psi = E\Psi$ la cual es una ecuación de autovalores, donde el operador actúa sobre una función (autofunción) que “devuelve” a la función multiplicada por un escalar (autovalor). En este tipo de operaciones se desea encontrar como solución la forma de la función (Ψ) tal que al ser operada por $\hat{H}$ provea como resultado la misma función Ψ, multiplicada por una constante (E en este caso). Por lo tanto, la mecánica cuántica establece que la energía E de un sistema, así como otras propiedades relacionadas, pueden obtenerse resolviendo la ecuación de Schrödinger.

Debido a que esta ecuación sólo es computable exactamente para sistemas físicos de dos partículas (como los átomos hidrogenoides), se realizan aproximaciones para lograr resultados en sistemas extensos. La dificultad para resolver sistemas de tres partículas (o más) tiene que ver con la imposibilidad de estudiar independientemente el comportamiento de una partícula sin tener en cuenta sus interacciones con las otras. Es decir que cualquier solución que pueda encontrarse para átomos polielectrónicos o moléculas es solo una aproximación a la solución verdadera. Los métodos de resolución pueden dividirse en *ab initio* (o “a primeros principios”) y en semiempíricos. Los primeros proponen soluciones a la ecuación de modo más riguroso y no requieren del uso de parámetros experimentales. Sin embargo, debido al costo computacional que demandan, su uso se restringe al cálculo de propiedades en pequeñas moléculas. Como alternativa aparecen los métodos semiempíricos, los cuales son menos precisos pues introducen mayores simplificaciones y aproximaciones a la ecuación a resolver, además de recurrir a datos experimentales para la estimación de valores utilizados en los cálculos. Esto reduce dramáticamente el tiempo de computo por lo que se utilizan para moléculas más grandes.

De la descripción de los métodos de cálculo desarrollada en este apartado puede deducirse que la elección de las metodologías a utilizar dependerá de que tipo de información se quiere calcular, así como del tamaño del sistema en estudio. Como regla general podemos mencionar que actualmente los estudios a nivel *ab initio* se encuentran limitados a moléculas que contienen decenas de átomos, los semiempíricos se utilizan para estructuras con cientos de átomos y los métodos de mecánica clásica son útiles para estudio de macromoléculas (miles de átomos).

En cuanto a la información que es posible obtener, la mecánica clásica es útil para calcular la energía asociada a una determinada conformación de una molécula, identificar conformaciones

más estables, estudiar el movimiento molecular, etc. Por otra parte, los cálculos a nivel cuántico son muy útiles para determinar cargas parciales sobre los átomos, potenciales electrostáticos, momentos dipolares, energías asociadas a los orbitales moleculares, energías de enlace, calores de formación, etc. Estos valores pueden considerarse descriptores de una molécula y podrían ser utilizados como variables independientes para la generación de relaciones estructura-actividad. Cabe aclarar que estos descriptores de origen clásico y cuántico pueden considerarse dentro de la categoría de descriptores 3D pues son dependientes de la conformación.

Referencias

Bellera C. L. (2014) Búsqueda racional de nuevos fármacos antichagásicos inhibidores de la cruzipaína. Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

Di Ianni, M. E. (2015). Topología molecular aplicada a la búsqueda de nuevos fármacos para el tratamiento de la epilepsia refractaria (Tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

Gago Bádenas F. Métodos computacionales de modelado molecular y diseño de fármacos. En: Monografía I Diseño de medicamentos [Internet]. Real Academia Nacional de Farmacia; 1994. p. 253-311. Disponible en: http://www.analesranf.com/index.php/mono/issue/view/96

Gantner M. E. (2016) Topología molecular aplicada al reconocimiento de sustratos de la Proteína de Resistencia del Cáncer de Mama (BCRP) (Tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

Graham L. Patrick. (1995) An Introduction to Medicinal Chemistry. 5th ed. Oxford University Press, New York. 1995. ISBN 0-19-855872-4.

CAPÍTULO 5
Métodos indirectos. Búsqueda racional de fármacos

Lucas N. Alberca y Alan Talevi

1. Cribado virtual

Las etapas tempranas del descubrimiento de fármacos implican la identificación de nuevos *hits* contra un blanco molecular específico y las sucesivas optimizaciones para mejorar las propiedades farmacológicas de dichos compuestos. Como ya se ha mencionado en el Capítulo 2, la industria farmacéutica recurre principalmente al Cribado Farmacológico de Alto Rendimiento (o *High Throughput Screening* - HTS*)* como estrategia para identificar en tiempos breves nuevos andamios (*scaffolds*) moleculares con la actividad deseada. Sin embargo, las tecnologías necesarias para la implementación de campañas de HTS requieren una inversión muy grande que solo las compañías farmacéuticas y algunos centros académicos de países desarrollados pueden afrontar. Al costo del equipamiento en sí debe sumarse, además, los altísimos costos operativos que deben afrontarse para su utilización y mantenimiento.

Una metodología alternativa, también mencionada previamente como más racional y económica, es el Cribado virtual (en adelante, CV), también conocido como *Screening Virtual* o *Screening in silico* o Tamizado Virtual. Algunos autores describen al CV como el "uso de la informática de alto rendimiento para analizar grandes bases de datos de compuestos químicos con el fin de identificar posibles candidatos a drogas*".* Otros autores amplían esta definición considerando al CV como un "conjunto de técnicas computacionales que permiten, a partir de representaciones de la estructura molecular de los compuestos químicos almacenados en grandes bases de datos, identificar compuestos potencialmente interesantes desde el punto de vista farmacológico*".* De manera más simple, podemos definir al CV como la utilización de modelos o algoritmos computacionales para predecir cuáles compuestos de una biblioteca de compuestos químicos podrían tener una cierta actividad deseada.

A continuación, se enumeran algunas de las características que hacen a la racionalidad de esta estrategia:

a) **Eficiencia en cuanto al costo y tiempo necesarios para su realización:** Actualmente existen una gran cantidad de bases de datos de compuestos químicos que se pueden descargar y manipular de manera gratuita y que son actualizadas/expandidas de manera periódica o continua. Entre ellas pueden mencionarse ChEMBL, que a la fecha compila más de 1,8 millones de

compuestos químicos y más de 15 millones de datos de actividad biológica; PubChem, con más de 96 millones de compuestos y más de 230 millones de datos de bioactividad y; DrugBank que contiene alrededor de 10000 compuestos utilizados terapéuticamente o que se encuentran actualmente cursando estudios clínicos, de los cuales cerca de 9000 son pequeñas moléculas tipo fármaco. Además, muchos de los programas que se utilizan para aplicar las técnicas de CV son de libre acceso o se pueden conseguir licencias académicas gratuitas como el programa Chimera y los diferentes paquetes de ChemAxon. También existen servidores online en los cuales se pueden realizar CV sin necesidad de descargar programas. Más aún, dependiendo de la técnica de CV a utilizar, es posible realizar el cribado con una simple computadora de escritorio. En este sentido, el CV supera ampliamente al HTS, que en cualquiera de los casos requiere tecnología de alto costo. Finalmente, puede mencionarse que, en tanto seleccionados mediante técnicas con fundamento teórico, los *hits* que emergen del CV tienen mayores probabilidades de demostrar actividad biológica contra el blanco de interés que un compuesto químico que hubiera sido elegido al azar, sin criterio racional.

b) **Carácter teórico:** No es necesario disponer de una muestra física de los compuestos químicos a analizar ya que el proceso se realiza en base a representaciones digitales de las moléculas. Además, se pueden considerar colecciones de compuestos que aún no han sido sintetizados de manera de sintetizar o adquirir sólo aquellos que presenten altas probabilidades de poseer la actividad de interés.

c) **Carácter bioético:** Las metodologías *in silico* reducen la cantidad de compuestos a evaluar tanto *in vitro* como *in vivo* (ciñéndose a los conceptos de reemplazo/reducción del uso de modelos animales) y aumenta las probabilidades de resultados positivos en esas instancias. El CV cumple con los principios internacionales para la investigación biomédica que involucra animales (*International Guiding Principles for Biomedical Research involving Animals*, CIOMS), que propone anteponer simulaciones computacionales y modelos *in vitro* a los estudios en animales de laboratorio siempre que sea posible.

En general, las estrategias para la búsqueda y diseño racional de fármacos asistida por computadora son clasificadas en dos categorías: Métodos indirectos (también llamados, métodos basados en el ligando), en los cuales no es necesario conocer la estructura tridimensional del blanco molecular y métodos directos (o métodos basados en la estructura), en los cuales es indispensable poseer un modelo tridimensional de la macromolécula. Subrayamos aquí, además, que cuando pensamos en diseño de fármacos siempre estamos considerando la eventual obtención de un compuesto químico novedoso que no hubiera sido sintetizado hasta el momento. En cambio, el CV, mayormente se enfoca en colecciones de compuestos que ya se conocen, que fueron previamente sintetizados o aislados a partir de fuentes naturales.

El presente capítulo se centra en los métodos indirectos para la búsqueda racional de fármacos.

2. Métodos indirectos

Existe una gran cantidad de blancos moleculares de los cuales no se encuentra disponible su estructura tridimensional y no es posible realizar un modelo tridimensional confiable mediante técnicas de modelado comparativo. En esos casos, la información estructural de uno o más ligandos activos en el blanco molecular seleccionado puede utilizarse para identificar cuáles son las características estructurales responsables de la actividad biológica de los mismos.

La hipótesis implícita de los métodos indirectos es que las moléculas similares exhibirán propiedades de unión similares con respecto a un dado blanco molecular, en tanto la actividad biológica dependerá de la estructura química del mismo.

Los métodos indirectos de CV pueden utilizar características obtenidas de distintos niveles de representación de las estructuras moleculares de los ligandos, habitualmente representaciones bidimensionales (2D) y/o tridimensionales (3D). En los siguientes puntos de este capítulo se describirán tres grandes grupos de métodos indirectos: los basados en la similitud molecular, los basados en el farmacóforo y, las metodologías basadas en descriptores moleculares (QSAR).

2.1 Métodos basados en la similitud molecular

La similitud molecular se enfoca en ciertas características estructurales de los compuestos tales como presencia, ausencia, o frecuencia de determinadas subestructuras químicas o determinados grupos funcionales, para determinar qué tan similares son dos moléculas entre sí. Para realizar esta comparación se requieren tres componentes básicos:

1- Una **representación de la molécula** cuyos componentes codifiquen las características químicas y/o moleculares relevantes de la misma.

2- Un **sistema de ponderación** (asignación cuantitativa de peso/importancia) de tales características. Es decir, se requiere de un esquema que permita determinar cuáles características tienen la misma importancia y cuales características son más relevantes.

3- Un **coeficiente de similitud** que transforme la información contenida en la representación estructural de la molécula a un valor numérico. En general, este número valor se encuentra entre 0 y 1, donde 0 indica que ninguna de las características estructurales comparadas está presente en ambas moléculas y 1 significa que por el método seleccionado se verifica la identidad completa de la molécula. Esto último no necesariamente implica que dos compuestos comparados sean efectivamente idénticos. Hay métodos de cuantificación de similitud que son “ciegos” a ciertas diferencias entre dos moléculas: por ejemplo, un método que caracterice los grupos funcionales como dadores o aceptores de enlaces de hidrógeno será incapaz de distinguir entre una amina primaria y un hidroxilo.

Una forma frecuente de comparar cuantitativamente dos moléculas se basa en la comparación de secuencias de bits (denominadas *molecular fingerprints* o huellas digitales moleculares)

de las estructuras a comparar. Típicamente, las huellas digitales moleculares se construyen asociando cada bit a la presencia o ausencia de una propiedad estructural, las cuales pueden contener información bidimensional o tridimensional.

En la Figura 5.1 se presenta un proceso simple de comparación de tres moléculas utilizando como coeficiente de similitud el coeficiente de Tanimoto en su forma binaria (no considera el número de veces que aparece una subestructura o característica molecular en una dada molécula, tan solo su ocurrencia o no ocurrencia). Este coeficiente es uno de los más utilizados para cuantificar similitud y se define por la Ecuación 5.1:

$$\boldsymbol{S} = \frac{\boldsymbol{c}}{\boldsymbol{a} + \boldsymbol{b} - \boldsymbol{c}}$$

Ecuación 5.1: Índice de Tanimoto.

Donde $\boldsymbol{a}$ y $\boldsymbol{b}$ representan el número total de subestructuras (o propiedades estructurales) presentes en cada una de las moléculas comparadas y, $\boldsymbol{c}$ indica el número de subestructuras comunes entre el par de moléculas a comparar. En el ejemplo de la **Figura 5.1** solo se consideraron ocho subestructuras para construir las huellas digitales moleculares de las moléculas A, B y C. La molécula A presenta seis de las ocho subestructuras consideradas, mientras que la molécula B presenta siete de estas subestructuras y la molécula C presenta sólo cinco. A partir de la comparación de a pares de las huellas digitales obtenidas, pueden calcularse los correspondientes coeficientes de Tanimoto. En el caso del ejemplo, los coeficientes asumen los valores S_{AB} = 0,625, S_{AC} = 0,571 y S_{BC} = 0,500, entonces, las estructuras A y B son las más parecidas de acuerdo con este método.

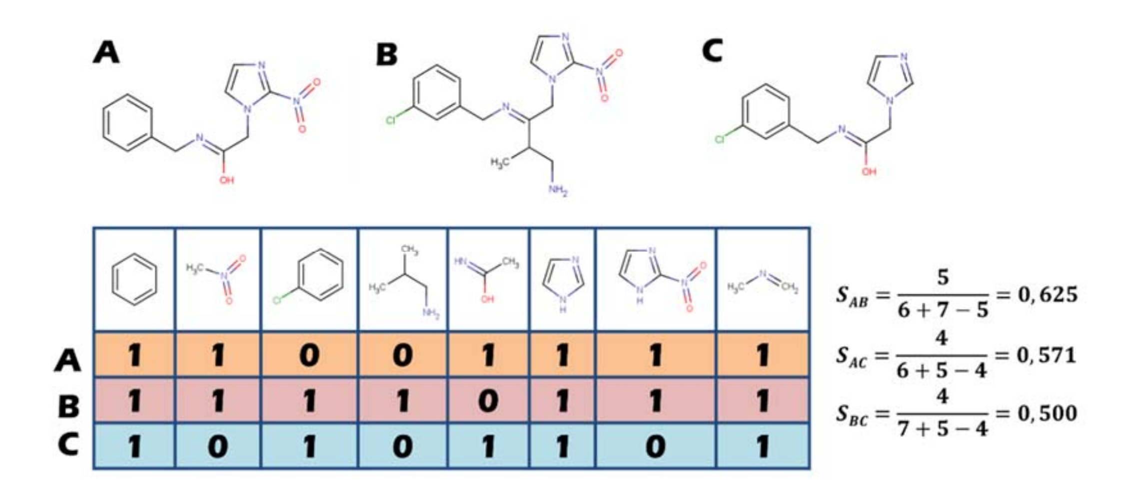


Figura 5.1: Esquema del proceso de comparación de tres moléculas por medio del coeficiente de Tanimoto.

Los métodos de similitud molecular son una buena estrategia para la realización de CV cuando el número de ligandos activos conocidos de un blanco molecular es limitado, o en el peor de los casos,

cuando se conoce un único ligando. Las técnicas de CV basado en la similitud consisten en cuantificar el grado de similitud entre la estructura que presenta la actividad deseada y las estructuras de los compuestos que integran la base sometida al CV. Una ventaja interesante de este tipo de técnicas es el bajo costo computacional que requieren, permitiendo realizar el cribado de grandes bases de datos en poco tiempo. Hoy en día, el uso de estas técnicas es casi trivial en tanto la mayoría de los repositorios químicos online utilizados para cribado incorporan funciones de búsqueda por similitud molecular. Por tal motivo, es raro actualmente el uso de estas técnicas de manera aislada, y se suele asociar a otros métodos indirectos o directos siempre que sea posible.

2.2 Métodos basados en el farmacóforo

El termino farmacóforo se refiere a un arreglo tridimensional de la molécula con las características moleculares mínimas para asegurar las interacciones con un blanco molecular especifico, logrando de esa forma que se produzca la respuesta biológica. Complementariamente, la IUPAC define a un farmacóforo como un conjunto de propiedades estéricas y electrónicas para asegurar las interacciones supramoleculares óptimas entre la molécula y un blanco biológico especifico activando o bloqueando su respuesta. En la Figura 5.2 se representa el farmacóforo que le da propiedades anestésicas a la cocaína, el cual fue utilizado para el posterior desarrollo del anestésico local procaína.

Figura 5.2: Ejemplo de farmacóforo (en azul) entre los anestésicos cocaína y procaína.

En este capítulo se presenta a los modelos farmacofóricos como aproximaciones basadas en el ligando, sin embargo, se debe tener en cuenta que este tipo de modelos cuentan con la versatilidad de poder ser abordados desde ambas perspectivas (receptor y ligando). La estrategia basada en el ligando consiste en superponer las estructuras tridimensionales de un conjunto de moléculas que comparten un blanco molecular y un sitio de unión (denominado conjunto de entrenamiento), con el objetivo de extraer aquellas propiedades químicas comunes que se presentan como esenciales para manifestar dicha actividad, en tanto que la perspectiva basada en el receptor propone postular los puntos de interacción necesarios en los ligandos y su disposición espacial en base a la complementariedad con las características conocidas del sitio activo. Un punto interesante para destacar es que

las características consideradas en un modelo farmacofórico no están exentas de un cierto grado de abstracción. Esencialmente, se consideran grupos aromáticos, lipofílicos, aceptores y dadores de hidrógeno, grupos ionizados positiva y negativamente. De este modo, una gran diversidad de funciones/grupos químicos puede cumplir con un dado requisito farmacofórico. Por ejemplo, un grupo amino primario o secundario o un grupo alcohólico pueden funcionar perfectamente como dadores de hidrógeno. Un sustituyente piridilo o un fenilo pueden, indistintamente, cumplir con el requisito de un anillo aromático en determinada posición espacial.

El modelado farmacofórico basado en el ligando es una opción computacional muy interesante (y, por cierto, gráfica y fácil de interpretar) cuando no se dispone de la estructura del blanco molecular. La generación de las hipótesis farmacofóricas basadas en múltiples ligandos comprende dos etapas principales:

1- Barrer el espacio conformacional para los ligandos del conjunto de entrenamiento de manera de representar la flexibilidad de cada uno de ellos.

2- Alinear las moléculas en aquella conformación que resulte en la superposición geométrica de la mayor cantidad de propiedades importantes para la actividad biológica.

La flexibilización conformacional de los ligandos y el alineamiento molecular son al mismo tiempo la base de la técnica y su principal dificultad. La etapa de alineamiento resulta ser la más costosa computacionalmente y la que representa un mayor desafío a la hora de diseñar algoritmos. Algunos métodos de alineamiento incorporan algoritmos donde el tiempo de cómputo crece exponencialmente con el número de propiedades analizadas, lo que limita la posibilidad de utilizar las hipótesis sobre grandes bases de datos y acota su campo de aplicación, con el equipamiento computacional actual, a moléculas pequeñas o a pequeñas bibliotecas. La etapa de búsqueda del modelo se facilita enormemente si se dispone de un análogo rígido activo, con escasa o nula libertad conformacional, que por lo tanto define la conformación activa o provee un fuerte indicio sobre la misma.

Las propiedades moleculares compartidas entre las moléculas del conjunto de entrenamiento que confieran la actividad (es decir, que estén implicadas en interacciones fundamentales con el blanco molecular) constituirán el farmacóforo. Si se desea utilizar este modelo para una campaña de CV, se busca en bibliotecas de compuestos químicos cuáles de ellos cumplen los requisitos establecidos. La desventaja de utilizar esta técnica para el CV es que todas las moléculas de la biblioteca a cribar requieren ser optimizadas conformacionalmente y superpuestas con el farmacóforo, un proceso que dependiendo de la cantidad de moléculas de la biblioteca puede ser muy costoso en términos de tiempo de cálculo.

Adicionalmente, los modelos farmacofóricos, a diferencia de otros métodos indirectos, tienen la ventaja de que pueden ser utilizados tanto para la búsqueda como para el diseño de nuevos fármacos. Una vez identificado el farmacóforo, es posible proponer modificaciones en un compuesto activo que permitan mejorar la actividad de los compuestos sobre el blanco molecular escogido.

2.3 Métodos basados en descriptores moleculares

Como se describió en el Capítulo 4, un descriptor molecular puede describirse como una variable numérica que representa algún aspecto o característica de la estructura molecular como el tamaño, el volumen, la cantidad de enlaces de hidrógeno disponibles, la lipofilicidad, la forma, la distribución electrónica, etc. En este punto se describirá cómo es posible encontrar relaciones entre una cierta actividad o propiedad de las moléculas y su estructura molecular codificada en descriptores moleculares.

2.3.1 Relaciones Cuantitativas Estructura-Actividad (QSAR)

Existen diferentes métodos estadísticos que permiten encontrar relaciones entre una variable dependiente (la propiedad o actividad que se desea modelar) y uno o más descriptores moleculares. Estas relaciones son conocidas como Relaciones Cuantitativas Estructura-Actividad o **QSAR** (del inglés, *Quantitative Structure-Activity Relationships*) y pueden ser descriptas por la siguiente Ecuación 5.2:

$$\boldsymbol{A = f(d_1, d_2, d_3, \ldots, d_d)}$$

Ecuación 5.2: Forma general de las ecuaciones QSAR.

Donde $\boldsymbol{A}$ es la actividad biológica (o una propiedad cualquiera) de un compuesto químico, que es considerada como una función matemática de ciertas características estructurales cuantificadas mediante descriptores moleculares (d_1, d_2). $\boldsymbol{A}$ puede pensarse como la variable dependiente, la "respuesta" (*response*) o la "salida" (*output*) del modelo.

La relación $\boldsymbol{f}$ puede ser obtenida por diferentes métodos estadísticos, métodos evolutivos, etc. Una vez obtenida esta relación, la actividad $\boldsymbol{A}$ de un compuesto nuevo o no testeado frente al blanco molecular seleccionado puede ser predicha a partir de su estructura molecular mediante el cálculo del valor que asumen los descriptores moleculares ($\boldsymbol{d_1, d_2}$, etc) (incluidos en el modelo QSAR) para ese compuesto.

2.3.2 Clasificación de los modelos QSAR

Existen diferentes formas de clasificación de los modelos QSAR:

a) Si se tiene en cuenta la ***naturaleza de la variable dependiente*** considerada, los modelos QSAR pueden clasificarse en:

Modelos QSAR cuantitativos: estos darán como resultado (o respuesta) una variable dependiente contínua, es decir, predecirán el valor numérico de la actividad (o propiedad) modelada. Por ejemplo, este tipo de modelos podría dar como resultado el valor de la constante de inhibición (Ki) de

un compuesto frente a un blanco molecular, la concentración del compuesto que genera un 50% de inhibición del blanco molecular (IC_{50}), el valor del logP de un compuesto, etc.

Modelos QSAR cualitativos (modelos clasificadores): darán como resultado una respuesta cuyo valor estará asociado a una categoría previamente establecida. Por ejemplo, las categorías pueden ser “inhibidores” y “no inhibidores” de un cierto blanco molecular, “tóxicos” y “no tóxicos” para células humanas, etc. Esta estrategia es muy útil cuando se entrenan los modelos con datos de diferentes fuentes (datos experimentales medidos en distintos laboratorios, por ejemplo) ya que un modelo clasificador permite amortiguar, al menos parcialmente, la variabilidad experimental presente en la base de datos.

b) De acuerdo a la **dimensionalidad de los descriptores moleculares** utilizados en los modelos, estos se pueden clasificar en:

Modelos QSAR de baja dimensionalidad: son aquellos modelos que utilizan descriptores moleculares independientes de la conformación. Por lo tanto, estos modelos se obtienen a partir de descriptores moleculares 0D, 1D y 2D. Los modelos constituidos enteramente por descriptores independientes de la conformación tienen la ventaja de que no requieren la optimización conformacional de las moléculas, por lo tanto, son poco demandantes computacionalmente. En los siguientes puntos de este capítulo se detallarán algunas estrategias útiles para la obtención de modelos de este tipo.

Modelos QSAR de alta dimensionalidad: son aquellos modelos que utilizan descriptores dependientes de la conformación de las moléculas. En esta categoría se agrupan los descriptores 3D y 4D. Las metodologías QSAR de alta dimensionalidad son más complejas y costosas computacionalmente respecto a las metodologías QSAR de baja dimensionalidad ya que requieren de la optimización conformacional 3D de cada una de las moléculas del conjunto de entrenamiento. La obtención de las conformaciones 3D de mínima energía de estas moléculas es el factor más importante para obtener modelos fiables ya que a partir de sus estructuras tridimensionales se calcularán los descriptores moleculares. La aplicación de modelos QSAR de alta dimensionalidad en CV requiere un tiempo considerablemente mayor que cuando se utilizan modelos QSAR de baja dimensionalidad ya que las bibliotecas de moléculas para el CV deben ser optimizadas conformacionalmente antes de la aplicación de los modelos. Los métodos QSAR-3D más populares, entre los que se destacan CoMFA, CoMSIA y GRID/GOLPE, se basan en colocar cada molécula optimizada en una grilla y calcular el valor que diferentes campos de fuerza (por ejemplo, coulómbico, estérico) ejercen sobre los vértices de la misma (colocando generalmente un átomo o molécula de prueba en cada vértice), estableciendo qué campo, y en qué zona de la grilla, es significativo para la actividad biológica.

c) Finalmente, si se tiene en cuenta en cuenta la manera en que se seleccionan las variables independientes y la forma en que se correlacionan con la actividad de interés, los modelos pueden ser de dos tipos:

Modelos QSAR Lineales: Se basan en la suposición de la existencia de una relación lineal entre las variables independientes y la variable dependiente o respuesta. Pueden ser cuantitativos, como cuando se generan mediante Regresión Lineal Múltiple (RLM), o cualitativos, cuando se obtienen mediante técnicas como el Análisis Lineal Discriminante (ALD).

Modelos NO Lineales: En estos tipos de modelos, los descriptores se asignan principalmente a un espacio relacional no lineal y ayudan a superar algunas limitaciones de los métodos lineales. Entre los métodos más utilizados para la construcción de modelos no lineales se encuentran los Arboles de Decisión, la Redes Neuronales Artificiales y las Maquinas de Soporte Vectorial para regresión (en inglés, *Support Vector Machines*).

2.3.2.1 Etapas involucradas en la construcción de modelos QSAR 2D clasificatorios (cualitativos) basados en descriptores[2]La construcción de ecuaciones QSAR, que relacionan la estructura de los compuestos químicos con una determinada propiedad biológica, conlleva un protocolo específico que implica varias etapas como puede apreciarse en la Figura 5.3.

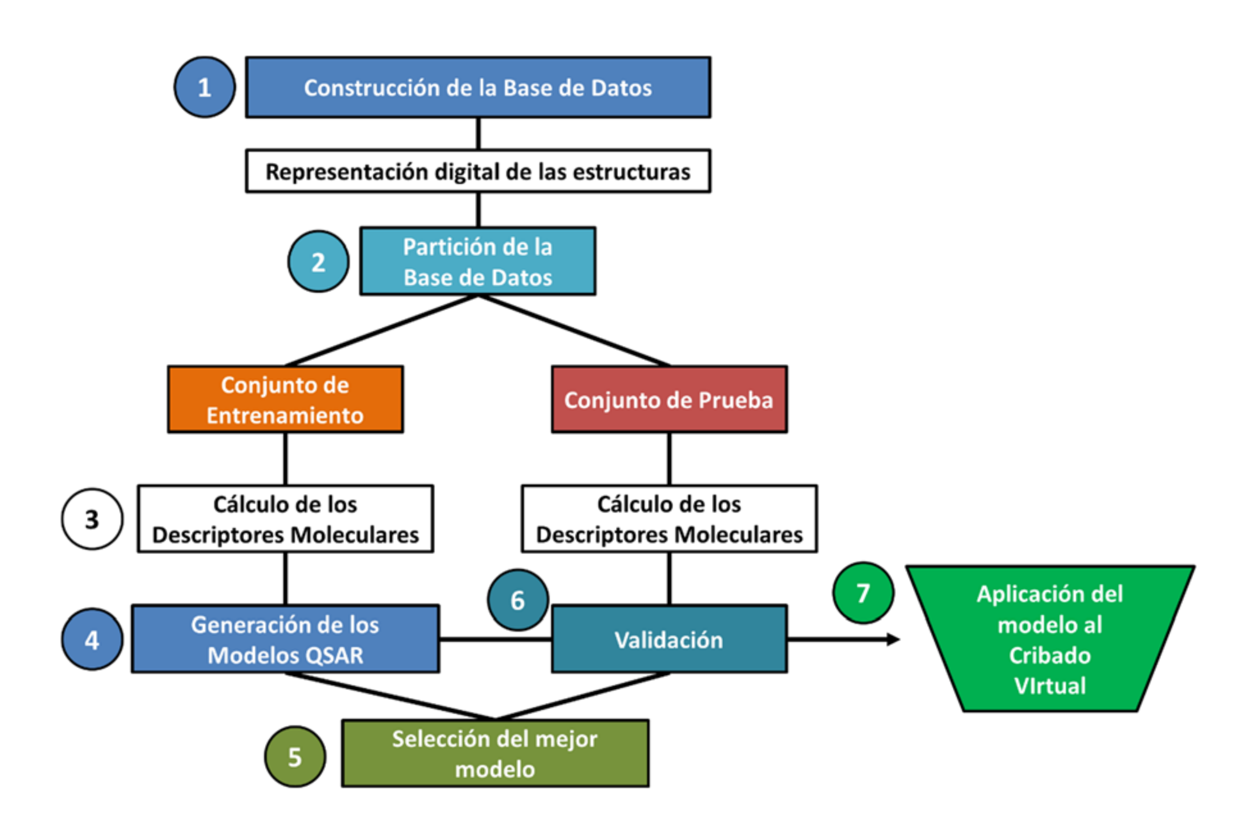


Figura 5.3: Esquema de las etapas vinculadas al desarrollo de un Modelo QSAR y su aplicación al CV.

[2] Las etapas generales para los modelos 3D son similares a las que se detallaran en este apartado, sólo que en todos los casos antes del cálculo se descriptores se agrega una instancia de búsqueda conformacional, y en algunos casos (por ej. CoMFA) una de alineamiento.

1) Compilado o construcción de la base de datos, set de datos o *dataset:*.

El primer paso para la obtención de modelos QSAR consiste en la búsqueda y recolección de compuestos químicos con sus correspondientes datos experimentales de la actividad o propiedad modelada. Con los datos obtenidos de esta búsqueda se genera un set de datos que posteriormente debe ser adecuadamente curado. Por ejemplo, si nuestro objetivo es generar un modelo que permita diferenciar entre inhibidores y no inhibidores de una enzima, entonces buscaremos en bibliografía y/o bases de datos como ChEMBL qué compuestos químicos han sido ensayados sobre esa enzima, cómo han sido evaluados y qué resultados se obtuvieron. Para generar modelos con buena capacidad de predicción es esencial realizar un curado de la base de datos que hemos obtenido; esto implica eliminar los compuestos duplicados, verificar la compatibilidad de los datos que provengan de diferentes tipos de ensayos de actividad, eliminar los compuestos con datos de actividad faltantes, etc.

El curado de la base de datos compilada es un paso crítico y esencial para la obtención de modelos QSAR fiables.

2) Partición de la base de datos:

Típicamente, para la construcción de modelos QSAR, la base de datos obtenida en el punto anterior es dividida en dos conjuntos:

Conjunto de entrenamiento: Conjunto de moléculas con las cuales se generarán/inferirán los modelos.

Conjunto de prueba: Conjunto de moléculas independiente del conjunto de entrenamiento con las que se validarán externamente los modelos (se estimará su capacidad de generalización o capacidad predictiva). Los compuestos del conjunto de prueba son “desconocidos” para los modelos ya que ellos fueron excluidos del proceso de desarrollo de los mismos.

Para realizar la partición de la base de datos en estos dos conjuntos se pueden utilizar diferentes estrategias. Uno de los métodos más utilizados es el de la **asignación al azar** de los compuestos del set de datos a cada uno de los conjuntos; sin embargo, con este método no se puede garantizar que los conjuntos de entrenamiento y de prueba sean representativos del conjunto de datos. Esto podría ocasionar que los compuestos del conjunto de prueba no sirvan para validar los modelos generados con el conjunto de entrenamiento, o viceversa, que los compuestos del conjunto de entrenamiento no produzcan modelos capaces de predecir acertadamente la variable dependiente para los compuestos del conjunto de prueba. Se ha demostrado que este tipo de partición es una aproximación adecuada cuando se seleccionan conjuntos de entrenamiento y de prueba de tamaño similar; en cambio, cuando los conjuntos de prueba son relativamente pequeños, se requieren aproximaciones de muestreo más racionales para proveer mejores resultados. Las mismas utilizan algoritmos de partición racional que intentan dividir la base de datos de una forma más representativa. Por ejemplo, existen numerosos métodos de agrupamiento (en inglés*, clustering*) racional basados en huellas digitales (*fingerprints*) y en descriptores

moleculares que pueden ser divididos en dos clases: los métodos jerárquicos como el agrupamiento de Ward y los métodos no jerárquicos como el de Jarvis-Patrick y k-means. La explicación de estos métodos excede el contenido de este libro, pero la idea fundamental es agrupar subconjuntos de compuestos con características moleculares comunes para luego muestrear representativamente de cada subconjunto una fracción de los mismos para el conjunto de entrenamiento y la fracción remanente para el conjunto de prueba.

Características del conjunto de entrenamiento: El conjunto de entrenamiento es una muestra de compuestos químicos de la cual se infiere una relación entre la estructura molecular y la actividad o propiedad de interés (el modelo QSAR). Se espera que esta relación sea generalizable a la "población" de compuestos químicos de la cual el conjunto de entrenamiento es representativo. Por lo tanto, la selección de los compuestos del conjunto de entrenamiento es crítica ya que define la región del espacio químico dentro de la cual la aplicación del modelo tendrá validez. El valor predicho por el modelo de la actividad/propiedad biológica modelada será confiable si y sólo si el compuesto predicho posee cierta similitud con alguna o varias de las estructuras incluidas en el conjunto de entrenamiento. El espacio químico dentro del cual la aplicación del modelo QSAR es confiable se denomina "dominio de aplicabilidad" o "dominio de aplicación" del modelo. Entonces, si se desea aplicar el modelo generado en campañas de CV sobre bibliotecas químicas que comprenden compuestos de gran heterogeneidad estructural, es conveniente que el conjunto de datos generado sea lo más abarcativo posible.

Otra característica que debe reunir el conjunto de entrenamiento es que los valores observados de la propiedad estudiada deberían presentar una buena distribución, idealmente uniforme en un rango de entre tres y cuatro órdenes logarítmicos de la propiedad modelada. El requisito de distribución uniforme es difícil de cumplir y se acepta como válida una distribución aproximadamente normal. La necesidad de una distribución uniforme tiene un fundamento lógico: por ejemplo, si tenemos un conjunto de entrenamiento formado mayormente por compuestos que poseen alta y mediana actividad, el modelo QSAR generado sólo va a predecir con exactitud valores de compuestos muy activos y moderadamente activos, o incluso estará sesgado hacia la asignación de altos valores de actividad, pero no nos permitirá diferenciar compuestos activos de inactivos. En otras palabras, el modelo no identificará qué características estructurales son desfavorables a la actividad por no estar los compuestos poco activos e inactivos representados dentro del conjunto de entrenamiento.

Un último requisito que se le pide tradicionalmente a los compuestos del conjunto de entrenamiento es que posean un mecanismo de acción común, esto es, interaccionen con el mismo receptor o blanco molecular y presenten el mismo sitio de unión. No obstante, se han modelado exitosamente en muchas oportunidades respuestas fenotípicas "ciegas" al mecanismo de acción considerado.

3) Cálculo de descriptores moleculares:

Una vez generados ambos conjuntos de trabajo, se procede al cálculo del valor que asumirán los descriptores moleculares para las moléculas del conjunto de entrenamiento. Como ya

se mencionó, dependiendo del tipo de los descriptores que se calculen podremos obtener modelos QSAR dependientes o independientes de la conformación. Sin embargo, hay que tener en cuenta que para el cálculo de los descriptores dependientes de la conformación es necesaria la optimización geométrica previa de las moléculas. Entre los softwares disponibles para el cálculo de descriptores moleculares podemos nombrar el Dragon, RDKIT, PaDEL y WEKA, entre muchos otros.

4) Generación del Modelo QSAR:

La generación de modelos QSAR implica dos etapas

a) La selección de descriptores adecuados para la predicción de la actividad o propiedad analizada.

b) La obtención del modelo matemático óptimo que correlacione la actividad o propiedad específica con los descriptores seleccionados.

Como ya se mencionó, los métodos de **selección de variables** (en este caso, descriptores moleculares) se presentan en dos grupos, los métodos lineales y los enfoques no lineales. En la Tabla 5.1 se enumeran algunos de los métodos más usados para la selección de variables tanto lineales como no lineales. En todos los métodos las variables se introducen en el modelo a través de una forma algorítmica y una función de aptitud (o de "fitness") o ciertos criterios de selección determinan qué variable debe ser retenida o eliminada del modelo.

Tabla 5.1: Ejemplos de métodos de selección de variables.

Métodos Lineales	**Métodos No Lineales**
Selección por pasos hacia adelante	Algoritmos genéticos
Eliminación por pasos hacia atrás	Programación evolutiva
Selección por pasos (combina los dos métodos anteriores)	Recocido simulado generalizado
Método de sustitución	Sistema de colonia de hormigas
Método de sustitución mejorado	Optimización de enjambre de partículas

La correlación entre la actividad o propiedad de interés y la estructura molecular se basará en métodos de regresión/clasificación que ponderan la contribución de cada una de las variables seleccionadas a la propiedad modelada. Estos métodos también pueden ser lineales y no lineales. La elección del tipo de método a emplear varía en cada caso y no suele ser evidente en primera instancia dado que ninguna técnica es consistentemente mejor que todas los demás. Por regla general, se prefiere ir de los métodos más sencillos a los más complejos.

Entre los métodos que se utilizan para la construcción de los modelos independientes después de la etapa de selección de variables o que incluyen una estrategia de selección de variables se pueden nombrar: Regresión Lineal Múltiple, Regresión de Componentes Principales, Mínimos Cuadrados Parciales, Redes Neuronales Artificiales, Maquinas de Soporte Vectorial, Árboles de Decisión, etc.

Se puede pensar que los errores en los que el modelo incurrirá al predecir la propiedad serán mayores o iguales a los errores presentes en los datos experimentales; por lo tanto, se desea disponer de datos experimentales sin errores significativos.

Cuando el conjunto de entrenamiento es compilado a partir de datos de literatura, se presenta una limitación para desarrollar un modelo QSAR con una variable dependiente continua (por ejemplo, IC_{50} o Ki), dado que los datos reportados no corresponden a las mismas condiciones experimentales (provienen de diferentes ensayos en distintos laboratorios). Para superar esta limitación se puede recurrir a los modelos QSAR clasificatorios lineales como los que se obtienen por la técnica conocida como **Análisis Lineal Discriminante (ALD)** o modelos QSAR clasificatorios no lineales como los obtenidos mediante **Árboles de Decisión**. Debido a su sencillez, en las próximas líneas se presenta una breve descripción de estos dos métodos, sin embargo, cabe la aclaración de que no son los únicos utilizados en para la generación de modelos QSAR cuantitativos.

Análisis Lineal Discriminante:

Esta técnica permite obtener un modelo clasificador lineal capaz de distinguir entre compuestos que presenten la actividad deseada (por ej, inhibidores de cierta enzima) de compuestos sin la actividad deseada (en el ejemplo, no inhibidores de la enzima elegida como blanco molecular).

ALD es un método de aprendizaje supervisado destinado a encontrar una combinación lineal de variables independientes (en nuestro caso, los descriptores moleculares) capaz de diferenciar entre objetos de dos o más categorías (en nuestro anterior, dos: ACTIVOS e INACTIVOS). La función discriminante (FD) obtenida corresponde a un plano en el espacio k-dimensional (siendo k el número de descriptores incluidos en el modelo) que, idealmente, deja a uno y otro lado los compuestos activos e inactivos. La Ecuación 5.3 representa en general a una FD:

$$\boldsymbol{FD} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \boldsymbol{a_i d_i}$$

Ecuación 5.3: Forma general de una función discriminante.

Siendo $\boldsymbol{a_i}$ el coeficiente de regresión asociado al descriptor molecular $\boldsymbol{d_i}$. Cada categoría de objeto (ACTIVOS o INACTIVOS) se asocia a un valor de una variable dependiente binaria arbi-

traria que funciona como etiqueta de las categorías. Por ejemplo, el valor 1 se asocia a los compuestos ACTIVOS y el -1 a los INACTIVOS. Dado que la función que se busca no predecirá una variable continua sino la categoría a la que pertenece cada elemento, el ALD y otras técnicas destinadas a encontrar modelos clasificadores pueden ser útiles para manejar datos ruidosos.

Árboles de Decisión

Los árboles de decisión son ampliamente utilizados debido a su capacidad predictiva, su facilidad de interpretación y su robustez para trabajar con datos ruidosos. Esta metodología permite obtener árboles de clasificación mediante un método jerárquico divisivo. Para construir un árbol se utilizan reglas de división del tipo: dado un elemento **o** a ser clasificado por una propiedad d, "si d **>** x entonces el elemento o pertenece a la clase **A** y si $d \leq x$ entonces pertenece a la clase **B**" (siendo d en este caso un descriptor y x el valor de corte para ese descriptor). Se procede a realizar una serie de divisiones binarias de los datos en subconjuntos (nodos internos), hasta llegar a un árbol maximal donde se reparten todas las observaciones en la hojas o nodos terminales (ver Figura 5.4). En cada nodo del árbol, el algoritmo elige el descriptor que más eficazmente divide el conjunto de entrenamiento en subconjuntos enriquecidos en una clase u otra. Para el caso de la clasificación, cuando ninguno de los descriptores proporciona una ganancia de información, se crea un nodo terminal (hoja) donde se asigna la clase que está más representada en el grupo de compuestos asignados al nodo.

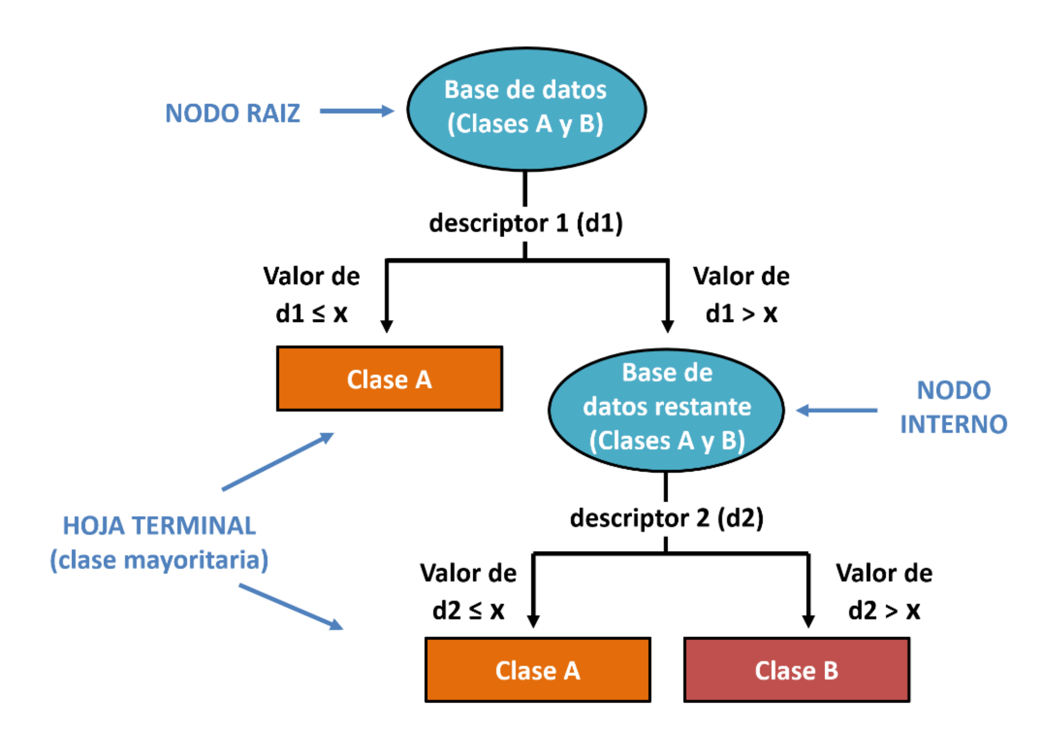


Figura 5.4: Árbol de decisión. "x" e "y" representan el valor de corte del descriptor de cada nodo.

5) Selección del Mejor modelo Generado:

Resultados de la validación externa: como se verá en el siguiente paso, la validación externa consiste en predecir la actividad de los compuestos del conjunto de prueba. Con la cantidad total de aciertos en la predicción, y la cantidad total de compuestos del conjunto de prueba se puede calcular el **% de buenas clasificaciones**, que es un buen estimador de la capacidad predictiva del modelo.

Principio de Parsimonia: también conocido como Navaja de Occam, es un principio metodológico y filosófico, según el cual "*No ha de presumirse la existencia de más cosas que las absolutamente necesarias*" es decir, cuando dos explicaciones logran explicar una misma observación, la más sencilla es la que se prefiere a priori. En este sentido, frente a la situación de encontrar varios modelos con buenos desempeños estadísticos, se optará por aquel que incluya menor número de descriptores (el más sencillo).

Sobreajuste (*overfitting*): se refiere al efecto de sobreentrenar un algoritmo de aprendizaje, es decir, que el modelo "memorice" las características del conjunto de entrenamiento (aumentando su poder explicativo o descriptivo) a expensas de sacrificar poder predictivo/capacidad de generalización (capacidad predictiva sobre compuestos no utilizados para el entrenamiento). Para los modelos lineales se exige una relación de al menos 10 entre el número de ejemplos de entrenamiento y el número de descriptores incorporados al modelo, de modo de reducir la probabilidad de sobreajuste.

En el caso de los árboles de decisión, se utilizan **algoritmos de poda** para evitar el sobreajuste. Estos remueven varias hojas y ramas del árbol, obteniéndose un árbol que tenga mejor poder de predicción como así también menor tamaño y complejidad. Básicamente hay 2 métodos de poda:

- post poda, que se lleva a cabo una vez que se finalizó la construcción del árbol
- poda en línea, que se realiza a medida que el árbol es inducido.

El principio que gobierna al proceso de poda es comparar la cantidad de errores que un árbol de decisión comete antes y después de cada posible procedimiento de poda, de modo de reducir al máximo ese error.

6) Validación del modelo:

Un proceso clave en el desarrollo de modelos QSAR es la validación de los mismos. Los procedimientos por los que se evalúan la robustez y capacidad predictiva de los modelos y, por lo tanto, su capacidad de predecir la actividad biológica (o la categoría) de compuestos aún no evaluados componen las técnicas de validación. Los métodos de validación se pueden clasificar en dos categorías: ***Validación interna***, que utilizan los compuestos del conjunto de entrenamiento para realizar la validación, y ***Validación externa***, que utiliza para evaluar la capacidad predictiva compuestos que no fueron utilizados para la generación de los modelos. Es de suma importancia utilizar ambos métodos para seleccionar modelos de buena calidad.

Validación interna

Validación cruzada (*cross-validation*): Las estrategias por excelencia son las técnicas denominadas *Leave One Out* (LOO, "Dejar uno afuera") y *Leave Group Out* (LGO, "Dejar un grupo afuera"). Estas estrategias permiten determinar la robustez de los modelos QSAR generados identificando si entre los compuestos del conjunto de entrenamiento hay algunos que influyen en mayor medida en el modelo generado (descriptores incluidos, valor de los coeficientes de regresión). La técnica consiste en la remoción de un número n (n igual a 1 corresponde a la técnica de LOO) de compuestos del conjunto de entrenamiento, la generación del modelo con los compuestos restantes, y la evaluación del nuevo modelo en el conjunto de compuestos removidos. Este proceso se repite hasta que todos los compuestos del conjunto de entrenamiento han sido removidos y evaluados por lo menos una vez.

Test de Aleatorización de Fisher: este ensayo se utiliza para descartar la probabilidad de correlación azarosa entre las variables independientes (los descriptores) y la variable dependiente (la propiedad o actividad). El método supone aleatorizar los valores de la variable dependiente en el conjunto de entrenamiento, cancelando por tanto cualquier relación que pudiera existir entre la estructura y la actividad. Posteriormente, se generan nuevos modelos para evaluar si existe probabilidad de correlación azarosa entre las variables independientes y la variable dependiente. Idealmente se espera que los modelos generados por aleatorización sean estadísticamente inferiores al modelo elegido.

Validación Externa: Esta validación consiste en evaluar la capacidad que poseen los modelos en la predicción de la actividad de un conjunto de moléculas que no hayan sido utilizadas para la generación de los modelos. Aunque las validaciones internas hayan obtenido buenos resultados, la gran mayoría de los especialistas coinciden en que el poder predictivo de los modelos QSAR sólo se puede establecer si el modelo se aplica con éxito para predecir los compuestos de un conjunto de prueba externo de tamaño adecuado. Para realizar esta validación, se aplican los modelos obtenidos en el conjunto de entrenamiento para predecir la actividad de los compuestos del conjunto de prueba.

Evaluación del desempeño de los modelos QSAR y selección del valor de corte.

Dos indicadores relevantes para estimar el desempeño de un modelo QSAR son su sensibilidad (Se, la tasa de verdaderos positivos) y su especificidad (Sp, la tasa de verdaderos negativos). Estos factores se definen por las Ecuaciones 5.4 y 5.5:

$$\boldsymbol{Se = \frac{VP}{VP+FN}}$$

Ecuación 5.4: Definición de Sensibilidad como indicador de desempeño de un modelo QSAR.

$$\boldsymbol{Sp = \frac{VN}{VN+FP}}$$

Ecuación 5.5: Definición de Especificidad como indicador de desempeño de un modelo QSAR.

Donde FN indica los falsos negativos (en nuestro caso, compuestos ACTIVOS que son clasificados por el modelo como INACTIVOS), FP indica los falsos positivos (compuestos INACTIVOS que son clasificados por el modelo como ACTIVOS), VN indica los verdaderos negativos (compuestos INACTIVOS clasificados como INACTIVOS) y VP indica los verdaderos positivos (compuestos ACTIVOS clasificados como ACTIVOS).

Cuando se utilizan modelos QSAR con fines clasificatorios cada molécula del conjunto de datos que se está evaluando obtendrá un resultado numérico único (o *score*) luego de la aplicación del modelo (volcado en el eje de las ordenadas, Figura 5.5). Se debe establecer un valor de corte del *score* a partir del cual se considerarán ACTIVOS e INACTIVOS los compuestos calculados. Por medio de la modificación de este valor de corte se puede observar que Se y Sp evolucionan de forma opuesta, por lo tanto, no es posible optimizar ambos parámetros de manera simultánea y se debe encontrar un balance adecuado entre los mismos. Para tal fin, se utilizan las curvas Receiver Operating Chracteristics (ROC), que son representaciones gráficas de la Se versus 1- Sp con las que es posible establecer el valor de corte óptimo de las funciones discriminantes (Figura 5.5).

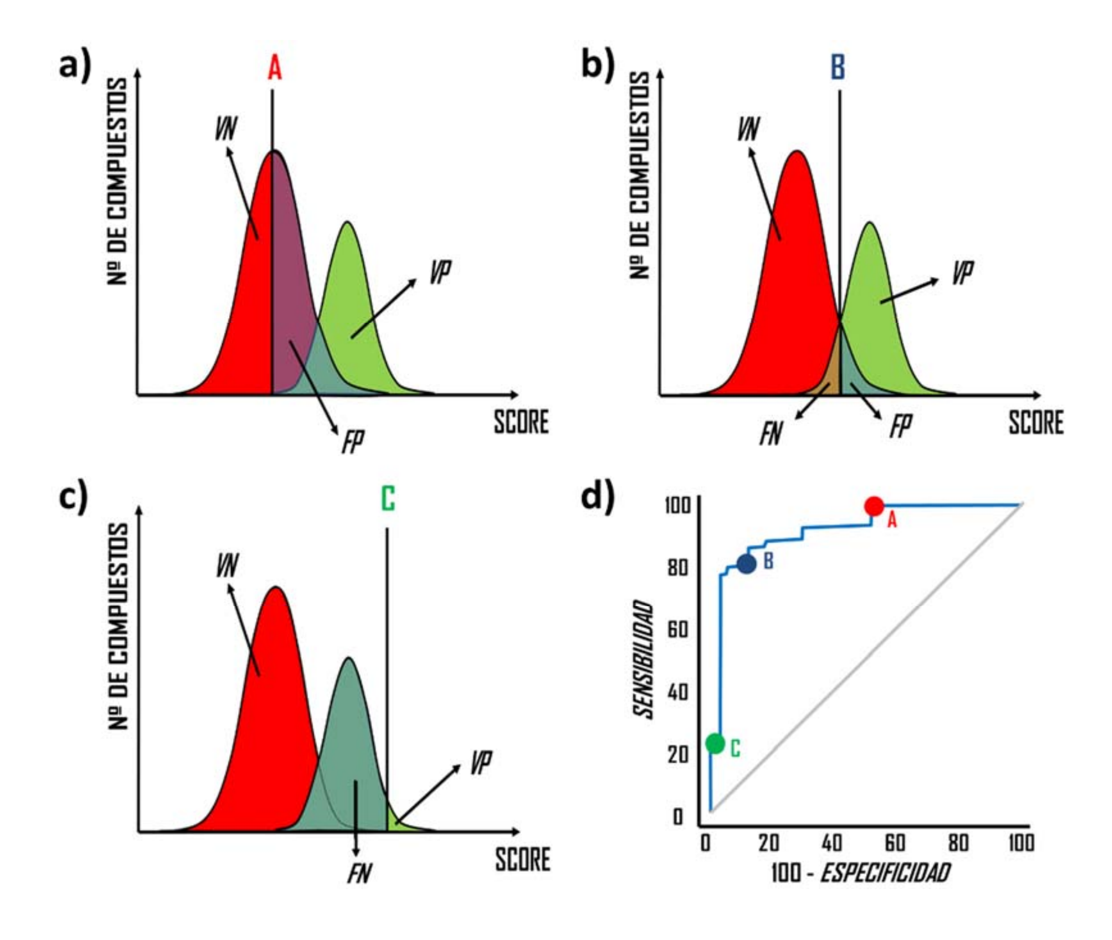


Figura 5.5: Esquema de la construcción de una curva ROC. Las figuras a), b) y c) representan gráficas de número de compuestos (eje y) versus el score obtenido por el modelo (eje x). Se puede observar como varían las proporciones de VN, VP, FN y FP cuando se modifica el valor de corte del score en los puntos A, B y C. d) Curva ROC.

Por otro lado, el **área bajo la curva ROC (AUCROC)** constituye una métrica valiosa para evaluar si el modelo se comporta significativamente mejor que una clasificación al azar y también permite comparar estadísticamente el desempeño de distintos modelos. Como puede observarse en la Figura 5.5, los ejes del gráfico de la curva ROC adoptan valores entre entre 0 y 100, aunque también pueden normalizarse a valores entre 0 y 1, por lo que un modelo ideal (Se y Sp igual a 1) corresponde a una curva que delimita una superficie cuadrada de área = 1. Un modelo clasificador que no difiera de la clasificación al azar obtendrá un AUCROC de 0,5. La capacidad de predicción de los modelos irá por lo tanto incrementándose a medida que el AUCROC aumente (se aleje hacia arriba de la línea diagonal) considerándose como un modelo clasificador perfecto a aquel que obtiene un AUCROC = 1.

La selección de un determinado balance entre Se y Sp no es una cuestión estadística, sino que depende del contexto. Por ejemplo, si pensamos en un laboratorio del ámbito público con bajos recursos económicos se priorizará Sp sobre Se para reducir el número de falsos positivos (compuestos que ensayarían en un test biológico, pero no tendrían la actividad predicha sobre el blanco molecular elegido). Esto significa, no obstante, que para reducir el número de falsos positivos se arriesga a perder andamios/motivos estructurales activos potencialmente valiosos. Lo contrario ocurrirá en laboratorios con altos recursos económicos, seguramente se priorizará la Se frente a la Sp de manera de no perder ningún candidato potencialmente valioso.

7) Cribado Virtual:

Una vez seleccionado y validado el mejor modelo QSAR se procede al CV de la base de datos elegida. Para llevar a cabo esta tarea, primero, para cada una de las moléculas de la base de datos a cribar, se realiza el cálculo de los descriptores moleculares que forman parte del modelo y luego, los valores obtenidos de los descriptores obtenidos se insertan en el modelo QSAR. De esa forma, el modelo obtiene un score (o una probabilidad) para cada uno de los compuestos de la biblioteca. Si el modelo generado era del tipo cuantitativo, el score representará el valor predicho de la actividad, mientras que si el modelo era del tipo QSAR cualitativo el valor del score se compara con el valor de corte seleccionado y de esa forma se predice la categoría o clase de cada compuesto.

Referencias

Bellera, C. L. (2014). *Búsqueda racional de nuevos fármacos antichagásicos inhibidores de la cruzipaína* (Tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

Di Ianni, M. E. (2014). *Topología molecular aplicada a la búsqueda de nuevos fármacos para el tratamiento de la epilepsia refractaria* (Tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

Gantner, M. E. (2016). *Topología molecular aplicada al reconocimiento de sustratos de la proteína de resistencia del cáncer de mama (BCRP)* (Tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

Graham L. Patrick. An Introduction to Medicinal Chemistry. 5th ed. Oxford University Press, New York. 1995. ISBN 0-19-855872-4.

Talevi, A. & Bruno-Blanch, L.E. (2009). Screening virtual: Una herramienta eficaz para el desarrollo de nuevos fármacos en Latinoamérica. *Latin American Journal of Pharmacy*, 28(1), 141–150.

Triballeau, N. et al. (2005). Virtual screening workflow development guided by the “receiver operating characteristic” curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *Journal of Medicinal Chemistry*, 48(7), 2534–2547.

Truchon, J.F. & Bayly, C.I. (2007). Evaluating virtual screening methods: Good and bad metrics for the “early recognition” problem. *Journal of Chemical Information and Modeling*, 47(2), 488 508.

CAPÍTULO 6
Métodos directos. Búsqueda y diseño racional de fármacos

Melisa E. Gantner, Pablo H. Palestro y Luciana Gavernet

Los métodos de descubrimiento de fármacos basados en la estructura del blanco molecular requieren justamente del conocimiento de la estructura tridimensional de dicha macromolécula (generalmente una proteína), para posteriormente predecir su capacidad de interacción con pequeñas moléculas candidatos a fármacos.

La hipótesis central de la aproximación directa es que la habilidad de un compuesto para ejercer un efecto biológico determinado depende de su capacidad para interactuar favorablemente en un sitio particular del blanco molecular. En base a esta hipótesis es posible obtener nuevos compuestos con actividad biológica mediante el análisis del sitio de unión en la proteína. El docking molecular, por ejemplo, es capaz de simular el acoplamiento proteína-fármaco desde el punto de vista estructural y energético, por lo que es una de las herramientas directas disponibles más utilizadas para el cribado virtual de bases de datos o liberías de compuestos químicos. Adicionalmente, la dinámica molecular provee una herramienta más precisa para el estudio de las propiedades estructurales y energéticas de los complejos, lo que la hace una metodología útil sobre todo para el estudio del origen de la actividad de un candidato y su optimización estructural. Estos dos métodos serán estudiados en el presente capítulo, aunque debe mencionarse que existen otras metodologías directas para el descubrimiento de compuestos activos como la construcción de farmacóforos basados en la estructura o el diseño *de novo* de moléculas activas que interaccionan con un determinado blanco molecular.

1. Estructuras 3D de los blancos moleculares

Actualmente, la información relacionada con la estructura tridimensional de los blancos moleculares generalmente se obtiene a partir de métodos experimentales, como la cristalografía por rayos X o la Resonancia Magnética Nuclear (RMN).

La base de datos Protein Data Bank (PDB), creada en 1977, alberga en la actualidad más de 140000 estructuras obtenidas por estos métodos experimentales, aunque muchas de ellas corresponden a diversas formas cristalinas de una misma macromolécula, con lo que el número de macroestructuras diferentes es menor. Muchas de estas macromoléculas juegan un rol vital en

vías metabólicas críticas y pueden funcionar como blancos terapéuticos. Además, el PDB alberga complejos de estructuras unidas a sus sustratos naturales o a fármacos, lo que proporciona una información muy valiosa respecto del origen de las interacciones intermoleculares que justifican la actividad biológica.

La cristalografía de rayos X en proteínas (o ácidos nucleicos) consiste en hacer incidir un haz de luz de longitud de onda correspondiente a los rayos X (del orden de las distancias interatómicas) sobre un conjunto ordenado de moléculas que conforman un cristal. Este cristal actúa como un amplificador para aumentar la señal producida por la difracción del haz de luz incidente sobre las proteínas. Dado que el patrón de difracción generado depende de la densidad electrónica alrededor de los átomos, esta información servirá para deducir la posición de los núcleos atómicos y, en consecuencia, permitirá la elucidación de la ubicación de los átomos y finalmente se obtendrá la estructura completa.

Las herramientas fisicomatemáticas para la resolución de la estructura están basadas en la teoría de la difracción de la luz y se utiliza como medida experimental de partida a las intensidades de las ondas difractadas. Este método resulta excelente en el caso de macroestructuras rígidas que pueden generar cristales ordenados. Sin embargo, en el caso de proteínas capaces de presentar regiones más flexibles resulta menos preciso, ya que la densidad electrónica para cada átomo se encuentra dispersa en un espacio mayor. Una medida de la precisión en la que se define la posición de los átomos en la estructura está determinada por la resolución. Este valor se relaciona directamente con la información experimental colectada, la cual depende de la calidad del cristal irradiado. En la actualidad estructuras con valores de resolución de aproximadamente 1Å se consideran muy ordenadas y es posible deducir fácilmente la ubicación de los átomos en el mapa de densidad electrónica. En el otro extremo se encuentran los experimentos de baja resolución (valores de 3Å o más) los cuales permiten sólo la observación del contorno de la cadena proteica y la posición de los átomos debe inferirse en mayor medida. Podemos definir como regla general que, si analizamos un conjunto de estructuras obtenidas por distintos experimentos de rayos X, aquellas con menor valor de resolución reportada generarán una mayor confianza en la ubicación de sus átomos.

En resumen, la cristalografía de rayos X es una de las técnicas más utilizadas para la elucidación de la estructura tridimensional de macromoléculas. Para su realización resulta imprescindible obtener cristales difractantes de la proteína, lo que representa la mayor dificultad del experimento. Debe mencionarse también que, al estar basada en la densidad electrónica, la cristalografía de rayos X sirve para la localización de átomos pesados, pero difícilmente detecta átomos de hidrógeno (el más liviano de los átomos, que posee solo un electrón). Además, es posible que en ciertos casos resulte complicado discernir entre elementos con números atómicos semejantes.

Otra técnica utilizada para la determinación de la estructura de macromoléculas es RMN. Este método se basa en las propiedades de ciertos átomos que presentan spin nuclear, como es el caso del hidrógeno cuyo valor de spin nuclear es ½ para su isótopo más abundante (^{1}H).

En presencia de un campo magnético externo, el momento magnético del núcleo de hidrógeno puede alinearse de dos maneras: a favor del campo externo (más estable) o en contra del campo.

Por lo tanto, para cada valor de intensidad de campo magnético externo es posible aplicar una determinada cantidad de energía para conseguir la inversión de la orientación. En otras palabras, cuando la energía necesaria para generar la transición del momento de spin nuclear (dependiente del campo magnético externo) coincide con la energía aplicada por una radiación externa, se produce una absorción y puede registrarse una señal. Este experimento es útil para la determinación de la estructura de una molécula porque la frecuencia de absorción del protón no es igual para todos ellos, sino que depende de su ambiente, esto es, la densidad electrónica que hay a su alrededor y la presencia de otros protones cercanos. Dado que una molécula orgánica generalmente posee una estructura con átomos de hidrógeno con distinto entorno, suele generarse un espectro de absorción con más de una señal. Cada una de estas señales representa a grupos de protones equivalentes que servirán para deducir la estructura molecular.

Los compuestos orgánicos poseen mayoritariamente átomos de carbono, hidrógeno, oxígeno y nitrógeno. Además del átomo de hidrógeno, los átomos de carbono y nitrógeno poseen isótopos con spin ½ (^{13}C ^{y} ^{15}N), pero son los isótopos menos abundantes. Es por ello que la espectroscopia más usada es la de protones, aunque se pueden usar las otras acumulando señales.

En cuanto a la aplicación de esta técnica en la elucidación de macroestructuras como las proteínas, los espectros resultan sumamente complejos debido a la gran cantidad de señales superpuestas que estos compuestos generan. Por este motivo, en estos casos se recurren a experimentos de RMN denominados multidimensionales. Estos métodos son capaces de disminuir la superposición de señales y generar más información respecto a su correlación dado que permiten inferir qué núcleos son cercanos a otros. Además de detectar los átomos de hidrógeno, una de las mayores ventajas del uso de RMN en la determinación de proteínas es que el método provee información acerca de las moléculas en solución, en oposición a los experimentos de rayos X que dan como resultado un promedio estático de la estructura. Esto resulta útil en el caso de proteínas flexibles ya que permite obtener información dinámica de las mismas. De hecho, como resultado del experimento se obtiene un conjunto de conformaciones posibles de la macromolécula. La mayor desventaja del método es que resulta difícil de utilizar para proteínas muy grandes (unos 40 kDa o más) debido en gran parte a la complejidad de los espectros.

Hasta aquí se han mencionado las dos metodologías experimentales más utilizadas para la elucidación de estructuras 3D de macromoléculas, entre las que pueden encontrarse blancos moleculares útiles para el diseño de fármacos. Cuando no se dispone de dichas estructuras, se pueden construir modelos tridimensionales teóricos utilizando técnicas de modelado por homología, modelado de comparación de plegamientos (o *threading*), o por métodos *Ab Initio*. Estas herramientas se inspiran en el trabajo realizado por Anfinsen y colaboradores, quienes observaron que una cadena aminoacídica de ribonucleasa desnaturalizada en solución alcanza el estado nativo en un tiempo razonable. Estos experimentos sentaron las bases para demostrar que: (1) una proteína puede ser identificada únicamente por su secuencia de aminoácidos; (2) esta secuencia es la responsable de codificar la funcionalidad *in vivo* de la proteína y (3) la secuencia de aminoácidos contiene la información necesaria para que la proteína adquiera su forma funcional, basándose exclusivamente en las interacciones intramoleculares y con el solvente que la rodea.

Cuando una proteína nueva se descubre, lo primero que se determina es su estructura primaria (o secuencia). El modelado por homología es una metodología basada en la comparación de la secuencia de la proteína cuya estructura se quiere modelar (secuencia problema) con otra (u otras) cuya estructura es conocida (secuencia molde). En esta técnica la proteína problema y la que actúa como molde se consideran homólogas.

Debe tenerse en cuenta que la homología es una inferencia o una conclusión respecto de la existencia de una relación ancestral entre la secuencia problema y la secuencia molde; obtenida a partir de la comparación de éstas. Para llegar a esta conclusión inicialmente se debe realizar un alineamiento entre ambas, que permita establecer su parecido. Los parámetros para considerar durante la comparación son similitud e identidad, los cuales derivan de la observación del alineamiento. La similitud de secuencia hace referencia al porcentaje de residuos alineados que tienen propiedades fisicoquímicas similares, por lo que pueden ser más fácilmente sustituibles entre ellos. La identidad de secuencia se refiere al porcentaje de aminoácidos idénticos alineados entre las dos secuencias. Cuando un alineamiento revela una similitud significativa entre un grupo de secuencias, estas pueden considerarse como homólogas. Si se conoce la estructura y función de una de estas secuencias “molde", esta información puede ser transferida a aquellas homólogas cuya arquitectura tridimensional no ha sido determinada aún.

El procedimiento general para el desarrollo del modelado por homología incluye las siguientes etapas: búsqueda e identificación en bases de datos de proteínas relacionadas al problema que puedan ser utilizadas como molde; realineamiento de secuencias entre las proteínas molde seleccionadas y la problema; construcción del modelo tridimensional basado en las regiones alineadas y en la predicción de coordenadas de átomos no alineados; refinamiento del modelo mediante minimización energética y evaluación del mismo (Figura 6.1).

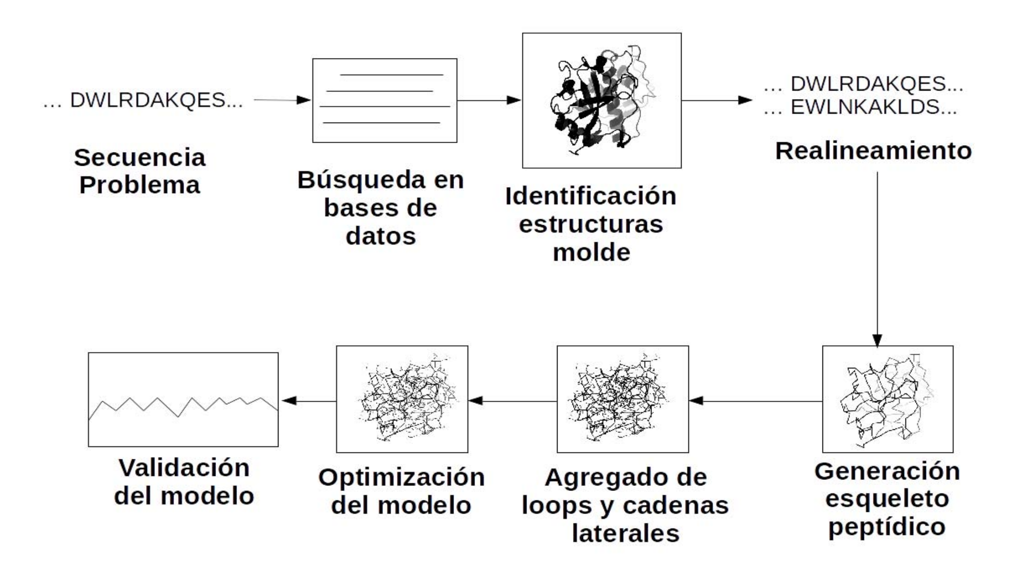


Figura 6.1: Esquema general de las etapas de un proceso de modelado por homología.

El método de comparación de plegamientos consiste en la asignación de las coordenadas del esqueleto carbonado de una estructura molde a la secuencia problema, y la medición de la plausibilidad del modelo tridimensional generado mediante el cálculo de una serie de parámetros empíricos. Mediante esta herramienta se pueden identificar proteínas con estructura conservada sin que presenten un origen evolutivo común. El fundamento de esta metodología es que, al momento, el número de plegamientos proteicos conocidos es muy reducido comparado con los millones de secuencias proteicas. En consecuencia, muchas proteínas pueden adoptar plegamientos similares incluso en ausencia de similitud en la secuencia. Esto admite el desarrollo de métodos computacionales para predecir la estructura proteica más allá de relaciones de homología.

Como su nombre lo sugiere, los métodos de predicción *Ab Initio* generan un modelo basado solo en la información obtenida a partir de la secuencia de aminoácidos de la proteína cuya estructura se quiere determinar. Los programas de predicción generalmente son diseñados utilizando los principios de la minimización energética: los algoritmos buscan las posibles conformaciones que puede adoptar la secuencia, para encontrar aquella que presenta un mínimo de energía. La ventaja que presenta este tipo de metodologías es que la predicción no está restringida a las estructuras tridimensionales conocidas para macromoléculas. Sin embargo, dado que las leyes fisicoquímicas que gobiernan el plegamiento proteico aún no son bien entendidas, las funciones de energía utilizadas al presente son inexactas.

2. Docking molecular

Como se mencionó anteriormente, la estructura tridimensional del blanco molecular (obtenida mediante métodos experimentales o por los modelos de predicción), resulta indispensable para el diseño de fármacos mediante métodos directos. Conseguida esta macromolécula, el paso siguiente es evaluar en forma virtual su capacidad de interacción con una molécula pequeña, posible candidato a fármaco. El docking molecular, o simplemente docking, es una simulación computacional orientada a predecir la conformación del complejo ligando-blanco, donde el blanco molecular es usualmente una proteína o un ácido nucleico y el ligando puede ser una molécula pequeña u otra macromolécula. El objetivo de esta herramienta es comprender y predecir el reconocimiento entre dos moléculas tanto a nivel estructural (encontrando el modo de unión entre ambos), como energético (prediciendo el valor de la afinidad de unión del complejo).

Al momento, las propiedades moleculares responsables de este reconocimiento son difíciles de identificar, entender y, sobre todo, de simular en una computadora. Sin embargo, los programas de docking pretenden encontrar la mejor forma de unión entre moléculas teniendo en cuenta diversos parámetros como la complementariedad geométrica, la flexibilidad de las estructuras y las interacciones intermoleculares. Como resultado estos software devuelven las orientaciones predichas del ligando en el sitio de unión del blanco molecular (también denominadas “poses").

Usualmente el proceso de búsqueda del modo de unión devuelve varios resultados posibles. Para determinar cuál es el más probable se utiliza una función de evaluación. Este valor ordena las poses obtenidas asociándoles un valor numérico relacionado con la energía libre de unión del complejo, que permite seleccionar las mejores posiciones relativas ligando-receptor (Figura 6.2).

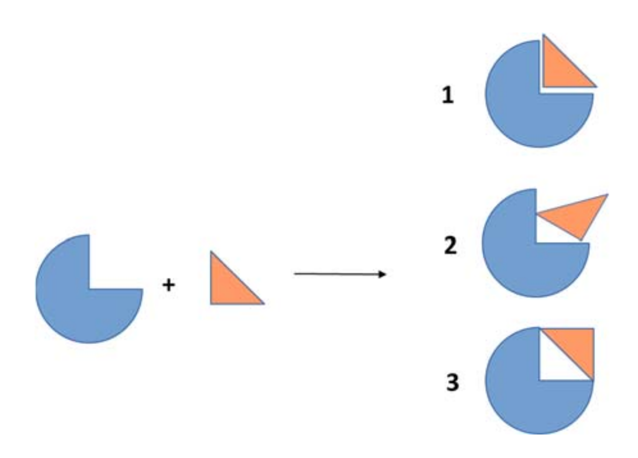


Figura 6.2: Esquema simplificado de una simulación por docking.

Por otra parte, existen diferentes aproximaciones respecto del grado de flexibilidad con que se considera el ligando y el blanco molecular en la simulación. Se habla de docking rígido cuando se propone al ligando y al blanco como cuerpos rígidos, limitando el proceso de búsqueda de poses a encontrar la mejor complementariedad geométrica/fisicoquímica posible. Por su parte, el docking flexible involucra el acceso a diferentes conformaciones del ligando, del blanco molecular o de ambos al mismo tiempo.

Las capacidades computacionales actuales permiten el empleo de docking flexible, aunque si se realiza un cribado virtual aparecerán limitaciones relacionadas con el elevado tiempo de cálculo que implica el docking de miles (o incluso millones) de moléculas sobre un blanco molecular. De hecho, considerar correctamente la flexibilidad del sistema es uno de los mayores desafíos en este tipo de herramientas de modelado.

De lo explicado anteriormente se desprende que en un proceso de simulación por docking conviven dos tipos de algoritmos: uno encargado de generar diferentes conformaciones relativas entre ligando-blanco molecular, y otro capaz de evaluar numéricamente las interacciones moleculares en cada una de esas conformaciones; a fin de determinar cuáles son las más probables. A continuación, se describen algunos ejemplos de estos algoritmos.

2.1 Métodos de búsqueda de conformaciones

El modo de unión del ligando con respecto al blanco molecular puede ser definido por la conformación adoptada por el ligando, su posición y orientación en el sitio activo. Cada una de estas variables describe un grado de libertad en un espacio de búsqueda multidimensional.

Un algoritmo de búsqueda riguroso debería muestrear todos los modos de unión posibles entre las dos moléculas involucradas en el docking. Sin embargo, esto es impráctico en términos computacionales debido al gran tamaño del espacio conformacional. Además, el problema se incrementa en un cribado virtual si se considera, por un lado, la flexibilidad del receptor y por otro el tamaño de las bases de datos de miles de ligandos que suelen utilizarse. Consecuentemente, sólo una pequeña parte del espacio conformacional suele ser muestreada, y esta exploración debe ser realizada con la suficiente exactitud como para identificar la conformación que mejor interacciona con la estructura del blanco molecular.

Los métodos de búsqueda más utilizados modifican la conformación del ligando y pueden clasificarse como sistemáticos o estocásticos. Debe mencionarse aquí que en el Capítulo 4 hemos detallado métodos de búsqueda de conformaciones en otro contexto (cálculo de descriptores de alta dimensionalidad), pero los algoritmos matemáticos utilizados son similares a los que se detallan en este capítulo.

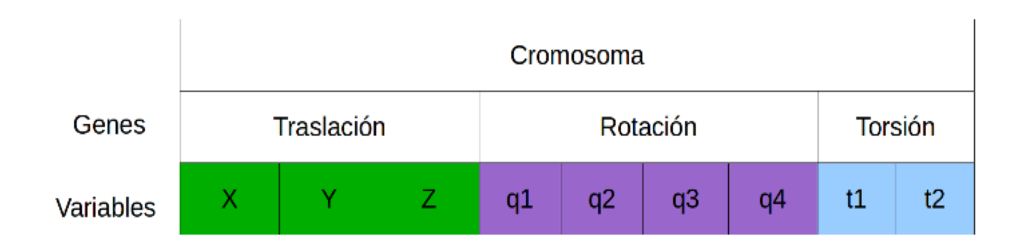
Entre los métodos sistemáticos se encuentran las metodologías de búsqueda de conformaciones generadas mediante la modificación sistemática de sus ángulos de torsión en un incremento fijo (por ejemplo, cada 10 grados). Se crea, entonces, un gran número de conformaciones definidas por todas las combinaciones posibles de valores de ángulos de torsión, las cuales son analizadas en su interacción con el blanco molecular. El mayor problema de esta técnica es el gran número de confórmeros creados, el cual se incrementa enormemente con el número de enlaces móviles. Por lo tanto, la aplicación de este tipo de métodos, en su forma pura, es muy limitado. Normalmente se aplican restricciones en los ángulos de torsión y limitaciones para reducir la dimensionalidad del problema.

Otro método sistemático muy utilizado es el de la fragmentación, el cual consiste en “cortar” el ligando para generar fragmentos de su estructura, los cuales son analizados en su interacción dentro del sitio activo mediante la función de evaluación del docking. Inicialmente se realiza el docking con el fragmento de mayor tamaño (o con aquellos que son rígidos) y sobre la mejor conformación obtenida se van añadiendo los siguientes fragmentos a la estructura (en todas las conformaciones posibles, quedando la más estable) de modo de ir construyendo gradual e incrementalmente el ligando completo dentro del sitio de unión del blanco molecular.

En cuanto a los métodos estocásticos, estos operan generando cambios al azar en los grados de libertad del ligando partiendo de una conformación inicial y, en algunos casos, también en residuos del blanco molecular. La nueva conformación obtenida es evaluada con la función de evaluación del docking, que analiza la energía de interacción en el complejo

formado. Con este resultado se examina si la conformación es aceptada o rechazada, mediante una función de probabilidad. Un ejemplo muy utilizado de este tipo de metodologías son los algoritmos genéticos.

Los algoritmos genéticos adaptan los conceptos de la teoría evolutiva a la resolución de problemas computacionales. En general, el proceso de búsqueda se basa en la generación de una población de soluciones candidatas sobre las que iteraciones sucesivas dan lugar a una selección competitiva, que elimina aquellas soluciones con un valor desfavorable de la función de evaluación. En la aplicación de los algoritmos genéticos al docking molecular, una pose particular de un ligando sobre su blanco puede ser definida por un conjunto de variables que describen la traslación, orientación y conformación de este sobre la macromolécula. Cada una de estas compone un gen, que en su conjunto constituyen el genotipo o cromosoma; mientras que la posición del ligando relativa al sitio activo de la proteína corresponde al fenotipo (Figura 6.3). Dicho de otro modo, una determinada conformación de un ligando en un blanco molecular puede codificarse en un cromosoma, el cual se divide en genes (representando cada una de las variables que describen la ubicación de los átomos del ligando). A su vez podemos agrupar a los genes en varios alelos, relacionados con los posibles valores que puede adoptar una variable concreta.



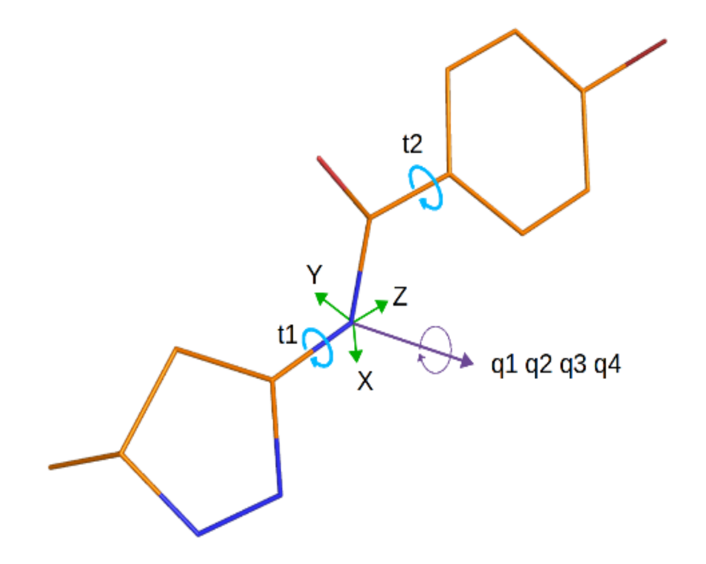


Figura 6.3: Esquema de los datos que definen un cromosoma utilizado en algoritmos genéticos para búsqueda conformacional. Incluye las coordenadas del centro de masa (X, Y y Z), los cuaterniones que definen la orientación y rotación del cuerpo rígido (q1, q2, q3 y q4) y los ángulos de torsión (t1 y t2) para la molécula de referencia.

El algoritmo inicia con la generación de una población de conformaciones de partida (cromosomas o individuos) al azar. Cada uno de los cromosomas son dockeados en el sitio activo del blanco y se seleccionan los mejores candidatos de acuerdo con la función de evaluación. Esto es, “sobreviven” los confórmeros que mejor interaccionan con el sitio activo y forman parte de la segunda generación de conformaciones, la cual se completa por la aplicación de operadores de recombinación y mutación sobre ellas. En el proceso de recombinación se produce el entrecruzamiento de genes entre un par de individuos seleccionados al azar, lo que genera nuevas conformaciones. Las mutaciones consisten en el cambio al azar de un alelo por otro en un gen, las cuales ocurren según una probabilidad predeterminada. Con el nuevo conjunto de candidatos obtenidos, la selección de los individuos para la nueva generación se realiza nuevamente por la función de evaluación del docking. El cálculo se finaliza luego de un número preestablecido de generaciones (Figura 6.4).

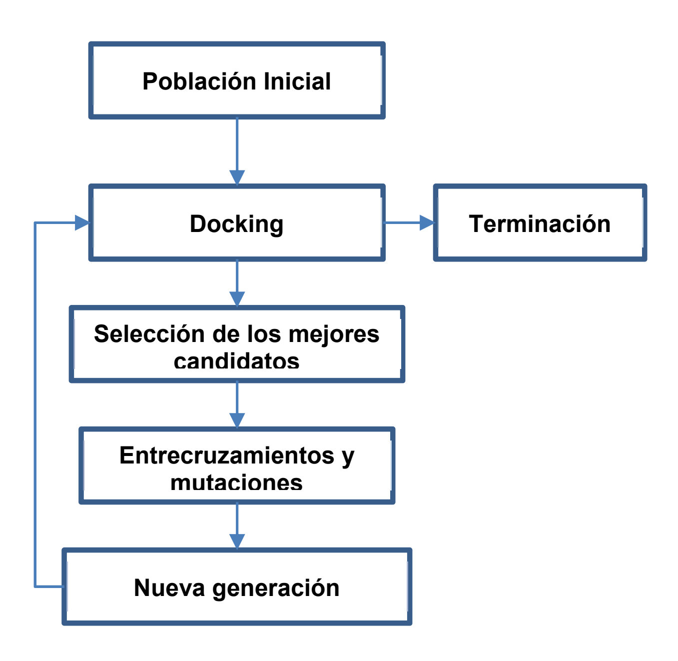


Figura 6.4: Esquema de búsqueda de conformaciones basada en algoritmos genéticos puros.

2.2 Función de evaluación

En una simulación por docking, la función de evaluación calcula en forma simplificada la energía de unión entre un ligando y un blanco molecular. Para ello utiliza cálculos basados en un tratamiento clásico de las moléculas que tienen como principal ventaja el reducido costo computacional respecto a los estudios mecano-cuánticos, por lo que resultan muy útiles en el estudio de sistemas donde intervienen macromoléculas.

Las funciones de evaluación aplicadas tradicionalmente se clasifican en tres grupos principales: basadas en campos de fuerza (*force field-based*), empíricas y basadas en el conocimiento

(*knowledge-based*). A continuación, detallaremos los fundamentos de aquellos basados en campos de fuerza por ser muy utilizados.

Como se mencionó en el Capítulo 4, los campos de fuerza se basan en la representación de la energía a partir de modelos clásicos y, por lo tanto, tienen como base el desglose de la energía potencial en componentes que describen las interacciones inter e intramoleculares que conformarán los distintos términos de la ecuación general. La descripción de los términos que contribuyen a la ecuación de energía de un sistema formado por una o más moléculas representan dos tipos de interacciones: las directas (generadas por átomos unidos entre sí o que interaccionan a través de uniones) y las que se presentan espacialmente, con la intervención de átomos ubicados en distintos puntos del sistema. Entre las primeras se encuentran aquellas relacionadas con uniones interatómicas, ángulos de enlace y ángulos de torsión; mientras que las segundas implican interacciones no enlazantes como aquellas del tipo van der Waals, puentes de hidrógeno o electrostáticas.

En cuanto a su forma funcional, las funciones de evaluación de docking se basan en los campos de fuerza diseñados para dinámica molecular (como se verá más adelante) y las interacciones ligando-blanco a menudo son descriptas mediante términos correspondientes a fuerzas del tipo Van der Waals y de Coulomb entre pares de átomos (Ecuación 6.1).

$$V = WvdW \sum_{i=1}^{N} \left(\frac{Aij}{rij^{12}} - \frac{Bij}{rij^{6}} \right) + WpH \sum_{i=1}^{N} E_{(t)} \left(\frac{Cij}{rij^{12}} - \frac{Dij}{rij^{10}} \right) + Welec \sum_{ij} \frac{qiqj}{4\pi\varepsilon_{(rij)} rij} + Wsol \sum_{ij} \left(S_i V_j + S_j V_i \right) e^{r_{ij}^{2} / 2\sigma^{2}}$$

Ecuación 6.1: Ejemplo de forma funcional para el cálculo de energía potencial. El primer término representa a la contribución de interacciones de Van der Waals, el segundo a las de puente de hidrógeno, el tercero a las fuerzas electrostáticas y el último a la desolvatación.

Las fuerzas de tipo Van der Waals son generalmente representadas mediante los potenciales de Lennard-Jones, dependientes de la distancia interatómica r_{ij}. Estos potenciales, también conocidos como potenciales 6-12 (en alusión a los exponentes numéricos) comprenden la suma de dos términos, uno atractivo que prevalece a distancias largas y otro repulsivo que, por el contrario, predomina a cortas distancias. La contribución electrostática está descripta en este caso como una formulación Coulómbica entre las cargas puntuales q_i y q_j, con una función dieléctrica dependiente de la distancia $\varepsilon_{(rij)}$, que reduce la contribución de las interacciones carga-carga. Algunas funciones también incorporan un término correspondiente a las interacciones de puente de hidrógeno, las cuales pueden representarse de distinta manera. Una de ellas determina la contribución de estas interacciones con un potencial similar al de Lennard-Jones, pero con exponentes 10-12 y un valor de $E_{(t)}$ que depende del ángulo t formado por los átomos que participan del puente de hidrógeno, ya que en este tipo de fuerzas influye la direccionalidad del enlace.

Debe mencionarse que el cálculo de la energía libre de unión ligando-blanco molecular implica también analizar la contribución entrópica del sistema. Esto resulta difícil de describir debido a la

ausencia de modelos matemáticos que la representen correctamente. Además, en el caso del docking molecular el solvente no se trata en forma explícita, lo que dificulta la modelización del fenómeno. Para subsanar en parte estas limitaciones, dentro de la función de energía potencial puede incluirse un término basado por ejemplo en el volumen de átomos (V_i) que rodean a un átomo dado y lo apantallan del solvente, ponderado por un término que mide la capacidad de solvatación (S) de cada átomo (el cual depende de la carga); y de una función exponencial que contiene una constante σ y depende de la distancia entre los átomos.

En cuanto a la pérdida de la entropía conformacional que implica la interacción del ligando con un blanco desde su estado no enlazado, es posible estimarla considerando que es directamente proporcional al número de ángulos de torsión (N_{tor}) del ligando (Ecuación 6.2).

$$\Delta S = W_{conf}\ N_{tor}$$

Ecuación 6.2: Algoritmo que representa la pérdida de entropía conformacional en una función de evaluación de docking.

Para el cálculo de todas estas ecuaciones es necesario especificar parámetros que reflejen al sistema real (por ejemplo A_{ij}, B_{ij}, $E_{(t)}$ etc.). Por tanto, dentro de la representación física de estas interacciones es fundamental la tipificación de los átomos del sistema. A cada uno de ellos debe asignarse lo que se llama "*AtomType*", que tendrá asociado los parámetros incluidos en los distintos términos de la ecuación general de energía potencial. En líneas generales *AtomType* además de contener información sobre un determinado átomo debe incluir el estado de hibridación de éste lo que definirá una geometría asociada a un potencial, como sp3 (tetraédrica), sp2 (trigonales), sp (lineales). Además, cada término contiene coeficientes optimizados mediante regresión lineal a partir de datos empíricos de energía libre de un conjunto de complejos ligando-proteína caracterizados experimentalmente (coeficientes W en las ecuaciones 6.1 y 6.2).

2.3 Validación de los programas de docking

Los distintos programas de docking combinan diferentes algoritmos de búsqueda y funciones de evaluación, resultando en distintas performances para distintos sistemas. Cuando se utilizan estos programas debe evaluarse su desempeño para el sistema en particular que nos interesa. Una manera de evaluarlos es utilizando un conjunto de testeo que nos permita validar la función de evaluación. Este conjunto incluye un subconjunto de compuestos que efectivamente interaccionan con el blanco molecular y cuya actividad es conocida, combinado con otro de estructuras que no presentan dicha actividad. Así, puede evaluarse la capacidad de los programas de docking de reconocer como activos (e inactivos) a aquellos que lo son dentro del conjunto de testeo. Una forma de cuantificar esta capacidad consiste en el uso de las curvas ROC del mismo modo que se utilizan en los métodos indirectos (detalladas en el Capítulo 5). En este caso la variable es la función de evaluación (energía libre de unión), por lo que estas curvas grafican la tasa de verdaderos positivos encontrados en función de la

tasa de falsos positivos, según varía el valor de corte. Al igual que para los métodos indirectos, a medida que el área bajo la curva ROC del modelo se acerca al valor 1 (clasificación perfecta), mayor será la capacidad discriminante del modelo seleccionado.

Otra herramienta de testeo de los programas es reproducir el modo de unión de un complejo conocido, lo que permite validar el algoritmo de búsqueda. Así, si se cuenta con la estructura experimental del complejo ligando-blanco molecular, una estrategia de validación es analizar la capacidad de imitar la conformación experimental de dicho complejo luego de la realización del docking con el programa seleccionado (re-docking).

Una vez validado el modelo, este puede utilizarse para predecir la interacción del blanco molecular en estudio con un candidato de interés o con bases de datos de pequeñas moléculas, donde el número de moléculas activas (y sus modos de unión) es desconocido. Esto último se denomina cribado virtual por métodos directos y será desarrollado en el apartado siguiente. A continuación, se muestra un esquema de las etapas a desarrollar en un proceso de docking de un ligando en un blanco molecular (Figura 6.5).

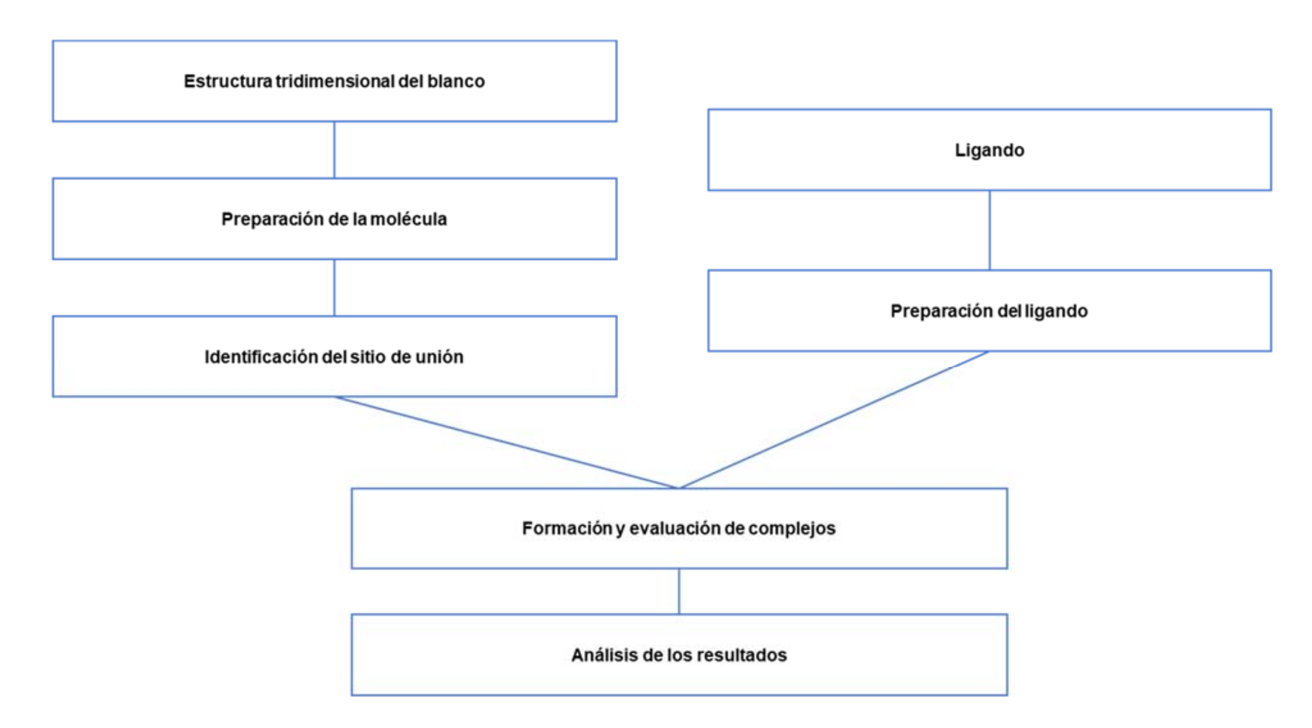


Figura 6.5: Esquema general de las etapas de un proceso de docking molecular.

La primera etapa consiste en disponer de la estructura tridimensional del blanco molecular, ya sea por haber sido elucidada por métodos experimentales o por haber generado un modelo computacional de dicha estructura. Es posible que en esta etapa la macromolécula deba ser acondicionada para los cálculos subsiguientes. Por ejemplo, si la estructura proviene de un experimento de difracción de rayos X, es posible que los átomos de hidrógeno deban agregarse a la estructura. También sucede frecuentemente que el sitio de unión se encuentra ocupado por otro ligando o moléculas de agua en la estructura experimental, las cuales deben eliminarse para “dejar espacio” disponible para realizar el docking del candidato a estudiar. Posteriormente debe identificarse el

sitio de unión. Esto generalmente es posible gracias a información experimental o a datos de literatura, aunque debe mencionarse que en ciertos casos el sitio de unión se desconoce y puede plantearse un docking ciego sobre toda la macroestructura o el uso de otros programas que buscan posibles sitios de unión dentro de una proteína. La siguiente etapa requiere preparar el potencial ligando, esto es, moléculas orgánicas pequeñas que deben estar debidamente protonadas al pH de trabajo y debe analizarse si existen posibles isómeros o tautómeros de las mismas que deban también dockearse. La tercera etapa, que es la parte medular del método, consiste en el docking propiamente dicho: el algoritmo (previamente validado) ubica al ligando dentro del sitio de unión en una gran cantidad de orientaciones/conformaciones y con su función de evaluación realiza un ranking de los posibles modos de unión del candidato. Finalmente se analizan los resultados y se evalúa la capacidad del ligando de interaccionar con el sitio activo.

3. Cribado virtual mediante métodos directos y mixtos. Ejemplos

Las etapas iniciales para realizar un cribado virtual mediante docking coinciden con las necesarias para realizar la simulación con una única molécula. Sin embargo, dado que el procedimiento devuelve una gran cantidad de posibles candidatos ordenados de acuerdo con su función de evaluación, resulta muy importante realizar un análisis previo de las capacidades de los programas para seleccionar compuestos que sean verdaderamente activos. Es decir, luego de la etapa de la validación del programa se debe encontrar un valor de corte (esto es, un valor de la función de evaluación límite) a partir del cual los compuestos se predicen activos. Este valor límite tendrá asociado un determinado porcentaje de sensibilidad y especificidad que satisfaga los requerimientos del investigador. Como ya se mencionó para el caso de métodos indirectos, la sensibilidad representa la tasa de compuestos activos pertenecientes al conjunto de testeo clasificados correctamente (es decir que superan el valor de corte) y la especificidad muestra la tasa de compuestos inactivos clasificados correctamente (que no superan el valor límite). Podemos resumir las etapas de un cribado virtual por docking como las siguientes (Figura 6.6):

1) Obtención de la estructura tridimensional del receptor. Como se discutió ampliamente al comienzo de este capítulo, la estructura puede estar determinada de forma experimental o alternativamente, se puede recurrir a alguna metodología de predicción como el modelado por homología.

2) Selección y validación del programa de docking que mejor se ajusta a nuestro sistema en estudio. Siempre que sea posible se debe validar tanto el algoritmo de búsqueda (re-docking) como la función de evaluación (conjunto de prueba o testeo). En esta etapa se realiza el cálculo de curva ROC y métricas de enriquecimiento del mismo modo que se utilizan en los métodos indirectos.

3) Selección del valor de corte óptimo de la función de evaluación para discriminar entre activos e inactivos según las preferencias del investigador (relación sensibilidad/especificidad adecuada). Una vez validado el modelo, ya puede utilizarse para realizar el cribado virtual, es decir,

predecir la interacción de nuestro blanco molecular con bases de datos de pequeñas moléculas donde el número de moléculas activas (y sus modos de unión) es desconocido.

4) Descarga de las representaciones digitales de las estructuras químicas de la base de datos a cribar y pre-tratamiento de todos los ligandos para definir el estado de protonación, los tautómeros e isómeros (en caso de poseer más de uno), cálculo de cargas soportadas por cada átomo y minimización energética.

5) Docking de todos los compuestos de la base de datos utilizando el modelo desarrollado y validado.

6) Ordenamiento de los compuestos según su valor de función de evaluación o score (de más negativo a más positivo).

7) Selección de los mejores candidatos (compuestos mejor rankeados con score más negativo) para su evaluación experimental. En este punto, además del score, se deben analizar las interacciones del compuesto en el sitio activo según la predicción del docking. Finalmente, la predicción del docking se combina con la experticia del investigador junto al conocimiento basado en la literatura para la selección final de los candidatos a evaluar experimentalmente.

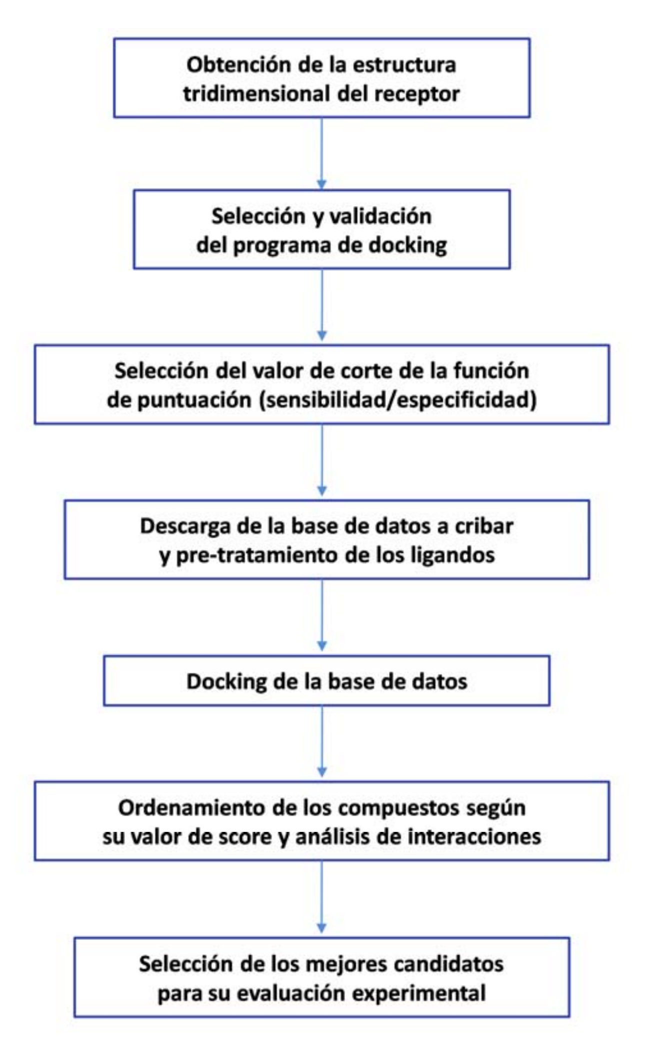


Figura 6.6: Esquema de las etapas del cribado virtual mediante docking molecular

Si bien las técnicas de docking se presentan como las más atractivas desde el punto de vista gráfico, las mismas muestran tres grandes limitaciones. La primera de ellas es de carácter coyuntural y tiene que ver con que la mayoría de las estructuras proteicas disponibles experimentalmente corresponden a proteínas solubles, dada la dificultad intrínseca ligada a la obtención de proteínas de membrana y transmembrana como canales iónicos, transportadores de eflujo, etc., en forma de cristales ordenados. La segunda limitación responde al costo computacional de la técnica comparada con los métodos indirectos, ya que es prácticamente imposible evaluar una gran cantidad de conformaciones de los ligandos cuando se trabaja con grandes bases de datos (como, por ejemplo, ZINC). En otras palabras, la metodología no permite analizar una gran cantidad de moléculas en un corto período de tiempo. Finalmente, la capacidad de predicción de los programas de docking es altamente dependiente del blanco molecular. La limitada exactitud de los programas de docking tiene su origen en las relativamente simples funciones de evaluación, que no logran reflejar completamente los fenómenos de desolvatación y reorganización del receptor y no capturan adecuadamente los factores entrópicos. Además, la mayor parte de las funciones de evaluación no consiguen modelar la unión de inhibidores covalentes ni iones. Todas estas limitaciones redundan en una alta tasa de falsos positivos.

Por otra parte, pese a que la utilización en cribado virtual de los métodos basados en la estructura y los métodos basados en ligandos individualmente han demostrado un inmenso potencial para recuperar hits iniciales, se hace evidente que muchas veces estos procedimientos no pueden cumplir todos los requisitos prácticos del descubrimiento de fármacos por sí solos. Además, con el aumento continuo del tamaño de las bibliotecas virtuales a cribar y el costo computacional asociado con algunos enfoques de cribado, especialmente del docking flexible, es indispensable integrar diferentes enfoques de cribado para filtrar compuestos. Los métodos mixtos, híbridos o combinados conjugan metodologías basadas en el ligando y basadas en la estructura de forma secuencial o paralela para realizar el cribado. Alternativamente se pueden combinar de forma secuencial o paralela métodos del mismo tipo. La forma más común de combinar estos métodos es usarlos en un embudo secuencial similar a lo que comúnmente se conoce como cribado virtual jerárquico o en cascada (HLVS, por sus siglas en inglés *hierarchical virtual screening*). Esta estrategia se basa en la aplicación secuencial de una serie de filtros computacionales para reducir una gran base de datos de pequeñas moléculas a un número de compuestos suficientemente pequeño (y promisorio) como para poder avanzar a los ensayos experimentales (Figura 6.7.A). En su mayoría, se utilizan enfoques basados en ligandos de bajo costo computacional, como la búsqueda por similitud, el screening farmacofórico o modelos QSAR durante los pasos iniciales de un protocolo HLVS. Los métodos que requieren recursos computacionales comparativamente altos, como el docking molecular e incluso la simulación de dinámica molecular, se utilizan una vez que el número de compuestos a analizar disminuye a un número razonable. La implementación de HLVS mixto es particularmente ventajoso en

los casos en los que se cuenta tanto con información de ligandos como información estructural, ya que toma en cuenta toda la información disponible en la identificación de nuevos hits, mejorando el rendimiento del cribado virtual al disminuir el número de falsos positivos y aumentar los verdaderos positivos, optimizando tiempos y recursos.

A diferencia del HLVS, en el cribado virtual paralelo (PVS) se ejecutan en paralelo varios métodos complementarios y los mejores hits rankeados según cada método se seleccionan para ser evaluados experimentalmente (Figura 6.7.B).

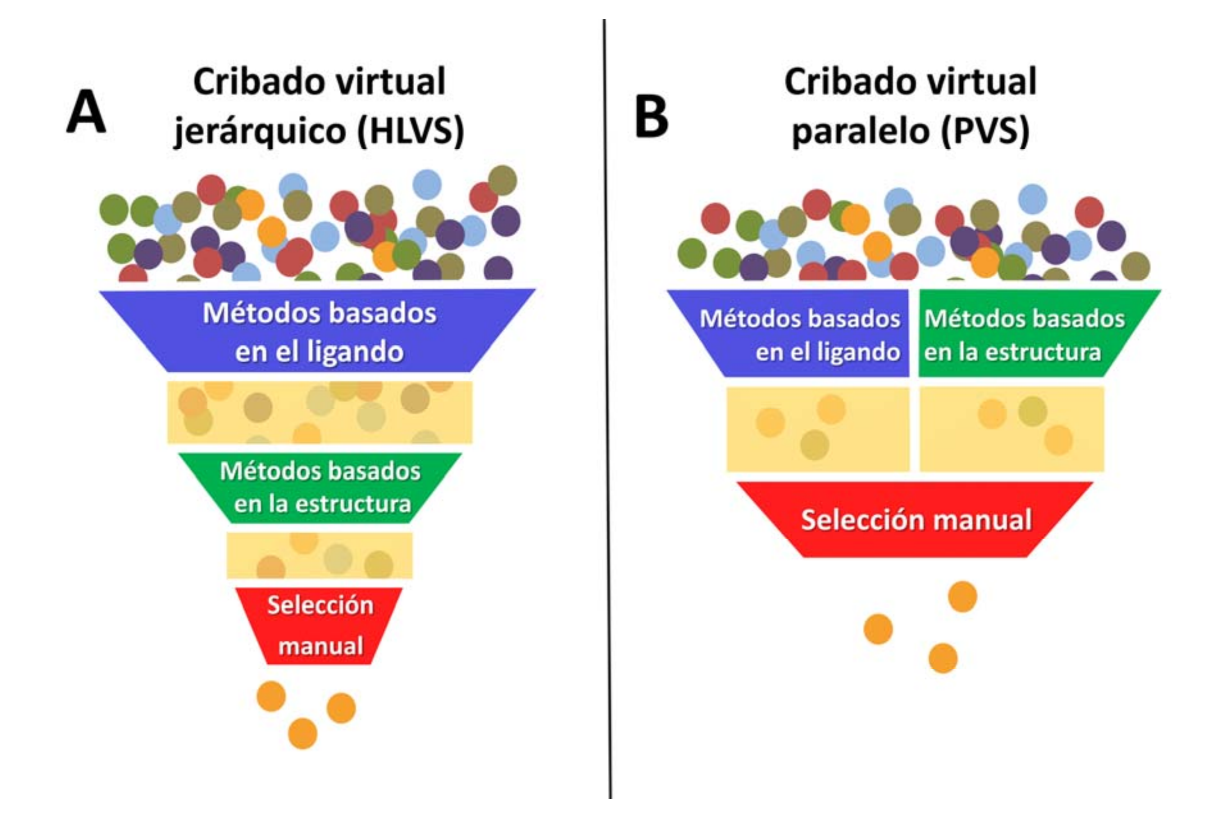


Figura 6.7: Cribado virtual mixto: integración de enfoques basados en el ligando y basados en la estructura. (A) Cribado virtual jerárquico (HLVS): una serie de filtros se aplican secuencialmente para reducir el número de compuestos que se seleccionarán para su evaluación experimental. Usualmente los métodos basados en ligandos, de menor costo computacional, se aplican en las primeras etapas, seguidos por los métodos basados en la estructura de mayor costo de cálculo. (B) Cribado virtual paralelo (PVS): los filtros basados en el ligando y en la estructura se aplican de forma independiente en el mismo (o similar) número de compuestos.

En la Tabla 6.1 se resumen algunos casos exitosos de descubrimiento de moléculas pequeñas mediante la aplicación de HLVS.

Tabla 6.1: Ejemplos de casos exitosos de moléculas pequeñas descubiertas mediante la aplicación de HLVS

Blanco molecular	**Acción**	**Estructura química del mejor compuesto encontrado**	**Actividad biológica**	**Métodos utilizados en el HLVS**
Transportador de serotonina	Neurotransmisión		Ki = 1.5 nM	fingerprints 2D, filtro ADME, modelado farmacofórico 3D y docking molecular
Proteína quinasa CK2	cáncer		85% de inhibición a 10 μM	Modelado bayesiano, modelado farmacofórico y docking molecular
Factor de coagulación VIII	Anticoagulante		IC50 = 3.5 μM	Modelado farmacofórico y docking molecular
Receptor del factor de crecimiento insulínico tipo 1 (IGF-1R)	Crecimiento celular, proliferación y apoptosis		IC50 = 57 nM	Modelado farmacofórico y docking molecular
ADN G-quadruplex	Envejecimiento celular y cáncer		Capacidad para unirse y estabilizar el G-quadruplex telomérico	Fingerprints 2D, búsqueda por similaridad y docking molecular

4. Dinámica molecular

La dinámica molecular clásica es una herramienta también basada en la mecánica molecular, la cual realiza la simulación de la interacción de átomos y moléculas a lo largo del tiempo. A diferencia del docking, en este tipo de simulaciones es posible calcular propiedades dinámicas, estructurales y termodinámicas del sistema en estudio. Para ello se genera información a nivel microscópico (como la posición y el momento de cada partícula) y mediante el uso de la mecánica estadística se la relaciona con observables a nivel macroscópico. En otras palabras, para conocer las propiedades dinámicas y termodinámicas globales de un sistema, resulta necesario conocer la probabilidad de hallar a dicho sistema en cada uno de sus estados microscópicos posibles. Debe entonces construirse un gran conjunto de réplicas del sistema para generar un agrupamiento de microestados (o ensamble) sobre el cual puedan calcularse las propiedades como un promedio. En la práctica, la construcción de este ensamble resulta imposible mediante simulaciones computacionales. Sin embargo, la dinámica molecular aporta los medios necesarios para poder estudiar la evolución temporal de un estado microscópico. Esto permitirá que el sistema pase por diferentes microestados accesibles y con ellos calcular un promedio de las propiedades en el tiempo. Esta idea de que un promedio temporal de una propiedad (a tiempo infinito) es igual al promedio del ensamble completo constituye la hipótesis ergódica de la mecánica estadística. Dado que no podrá lograrse una simulación a tiempo infinito, el muestreo de microestados obtenido por dinámica molecular será incompleto, y esto debe tenerse en cuenta sobre todo cuando se deseen calcular en forma muy precisa propiedades que dependan de estados de alta energía relativa.

La dinámica molecular genera las configuraciones de un sistema mediante la resolución de las ecuaciones de movimiento. Mediante la integración de la ecuación que representa la segunda ley de Newton (Ecuación 6.3) puede calcularse la trayectoria, es decir, la variación a través del tiempo de la posición y la velocidad de las partículas. Para esto es necesario conocer la fuerza que actúa sobre cada partícula, la cual en el caso de un sistema complejo como una macromolécula varía cada vez que estas cambian de posición.

$$\frac{d^2x_i}{dt^2} = \frac{F_{x_i}}{m_i}$$

Ecuación 6.3: Ecuación que representa la segunda ley de Newton. X*i* es la coordenada de la partícula i, *t* es el tiempo, *Fxi* es la fuerza sobre la partícula en la coordenada *xi* y *mi* es la masa de la partícula.

La Figura 6.8 muestra un esquema simplificado de un cálculo de dinámica molecular para un sistema de partículas. Con esta herramienta puede estudiarse, por ejemplo, un complejo formado por un blanco molecular y un posible candidato ubicado en el sitio activo para predecir sus interacciones y movimientos relativos.

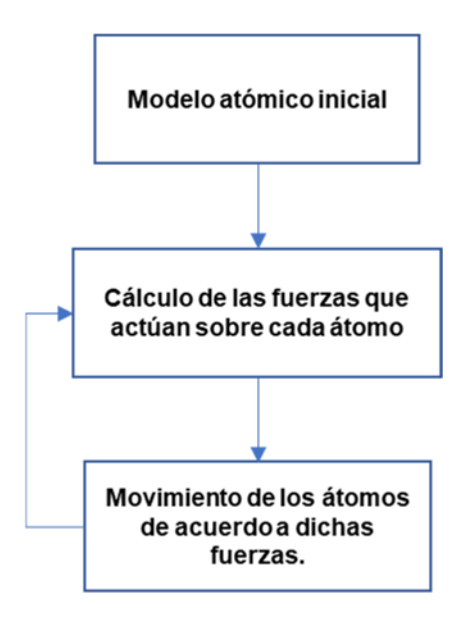


Figura 6.8: Esquema del cálculo de las posiciones de los átomos mediante dinámica molecular.

Inicialmente es necesario un modelo atómico del sistema a estudiar. En el caso de un complejo blanco-ligando, este puede ser extraído del PDB, el cual contiene estructuras de macromoléculas obtenidas experimentalmente. Otra alternativa es proponer como estructura de partida un modelo del blanco molecular mediante métodos computacionales según se ha detallado previamente (modelado por homología o threading). En el caso de desconocerse la posición relativa del ligando en el sitio activo puede realizarse un docking, de modo de proponer como estructura inicial del complejo aquel generado por dicha simulación. Con este modelo inicial el programa calcula las fuerzas que actúan sobre cada átomo, y se produce el movimiento de éstos de acuerdo con la segunda ley de Newton.

Las fuerzas sobre cada partícula se calculan mediante la derivada negativa (respecto de la posición) de la función energía potencial, cuya forma funcional está definida en el campo de fuerza del programa. La Ecuación 6.4 muestra un ejemplo de campo de fuerza de dinámica molecular. La energía potencial es función de la posición de los átomos y se calcula como una suma de energías internas o de enlaces y no enlazantes. Entre los primeros tres términos aparecen las contribuciones relacionadas con los movimientos en enlaces, ángulos y ángulos de torsión de las moléculas respectivamente; mientras que entre las últimas están los aportes de interacciones entre átomos no unidos entre sí como las del tipo coulómbicas o de van der Waals.

$$V = \sum_{enlaces} k_r\,(r-r_0)^2 + \sum_{ángulos} k_\theta\,(\theta-\theta_0)^2 + \sum_{enlaces} \frac{V_n}{2}[1+\cos(n\omega-\gamma)] + \sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{qiqj}{4\pi\varepsilon_{(0)}rij} + \sum_{i=1}^{N}\sum_{j=i+1}^{N}\left(\frac{Aij}{rij^{12}}-\frac{Bij}{rij^{6}}\right)$$

Ecuación 6.4: Ejemplo de forma funcional de un campo de fuerza utilizado en la dinámica molecular.

Tal como en el caso del docking molecular, la forma funcional de los campos de fuerza contiene parámetros cuyos valores varían según los tipos de átomos y usualmente son obtenidos utilizando datos experimentales o mediante cálculos mecano-cuánticos. Dado que en estos sistemas las fuerzas sobre cada átomo se modifican con el cambio de posición de la partícula (o de sus vecinas) existe un acoplamiento entre los movimientos que impide una resolución analítica de la ecuación de Newton (Ecuación 6.3). En estas circunstancias las ecuaciones de movimiento se resuelven usando el método de diferencias finitas, cuya idea principal consiste en partir la integración en pequeños intervalos de tiempo fijo δt, dentro de los cuales las fuerzas se asumen constantes. Una vez calculada la fuerza en este segmento es posible determinar la aceleración de las partículas y mediante el conocimiento de las posiciones y velocidades a un tiempo t es posible calcular las posiciones y velocidades a tiempo $t + \delta t$. Esto implica que la simulación "avanza en el tiempo" y el proceso puede repetirse un número predefinido de ciclos para generar el ensamble de conformaciones (Figura 6.8).

Resulta útil mencionar aquí que en los sistemas biológicos raras veces las macromoléculas se encuentran aisladas y suelen estar inmersas en un entorno de membrana o de moléculas de aguas. A diferencia del docking, la dinámica molecular permite analizar la influencia del entorno mediante el uso de moléculas de solvente en forma explícita. Por ejemplo, es muy probable que en un entorno acuoso exista un apantallamiento de las interacciones electrostáticas entre dos átomos expuestos al solvente; o que moléculas de agua sean partícipes de interacciones del tipo puente de hidrógeno con el ligando o el blanco molecular. Estas consideraciones hacen de la dinámica molecular una herramienta más precisa a la hora de estudiar la estructura, la dinámica y la termodinámica de sistemas biológicos.

Adicionalmente, la dinámica molecular permite analizar a los blancos moleculares como estructuras flexibles, a diferencia del docking molecular donde dicha flexibilidad no se tiene en cuenta o se considera parcialmente. Es por eso que en la actualidad el docking molecular es recomendable como herramienta para el cribado virtual de grandes bases de datos (ya sea como metodología única o como parte de un protocolo de cribado mixto), pero debe recurrirse a la dinámica molecular para la optimización de fármacos, pues esta última metodología permite un estudio mucho más profundo de las interacciones ligando-blanco molecular responsables de la acción biológica.

Referencias

Anfinsen, C. B. (1973). "Principles that govern the folding of protein chains". *Science*, 181, pp. 223-230.

Di Ianni, M. E. (2015). Topología molecular aplicada a la búsqueda de nuevos fármacos para el tratamiento de la epilepsia refractaria (Tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

Kitchen, D. B., Decornez, H., Furr, J. R., Bajorath, J. (2004) “Docking and scoring in virtual screening for drug discovery: methods and applications". Nature reviews. Drug discovery, 3, pp. 935-949.

Kumar A., Zhang K. Y. (2015). “Hierarchical virtual screening approaches in small molecule drug discovery”. *Methods*. 71, pp 26-37.

Palestro, P. H. (2016). Búsqueda de nuevos anticonvulsivos a partir de métodos computacionales (Tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

CAPÍTULO 7
Profármacos

María L. Villalba y Melisa E. Gantner

Como hemos mencionado al comienzo de este libro, el descubrimiento y desarrollo de un nuevo fármaco es un proceso que requiere, en promedio, entre 12 y 15 años de trabajo. Además de sus altos costos, existe una alta probabilidad de no alcanzar un resultado exitoso. Usualmente, a pesar de contar con un compuesto altamente activo que interacciona adecuadamente en su sitio de acción, la alta tasa de fracaso se debe a características fisicoquímicas intrínsecas del compuesto activo, las cuales derivan en problemas de tipo ADME, entre los que podemos mencionar problemas de solubilidad y/o de permeabilidad, alta tasa de metabolismo presistémico, toxicidad elevada, etc.

Una de las formas de abordar este tipo de problemas son las estrategias de optimización de fármacos, las cuales hemos detallado en capítulos anteriores. Otra de las estrategias es la síntesis de análogos del compuesto. Sin embargo, en algunas ocasiones la molécula presenta elementos estructurales tan restringidos que no pueden prepararse análogos mediante modificaciones habitualmente empleadas en química medicinal, sin comprometer su afinidad por el receptor o el blanco molecular. Debido a esto el diseño de profármacos adquiere relevancia, ya que permite modificar propiedades del fármaco de forma temporal, mejorando su biodisponibilidad y, por ende, su acción en sitios específicos. Como hemos mencionado previamente, un profármaco (o prodroga) es aquel compuesto que se administra en forma inactiva y que ejerce su efecto luego de ser metabolizado por el organismo.

En el diseño de profármacos una etapa clave es la incorporación de un mecanismo de activación que convierta el mismo en sus especies activas de manera eficiente y/o controlada hasta alcanzar el objetivo deseado (blanco molecular). Dicha activación puede ser promovida por enzimas hidrolíticas o por procesos de oxido-reducción.

Actualmente, entre un 8 y un 10% de los fármacos aprobados en todo el mundo son profármacos, y aproximadamente un 15% de los nuevos fármacos aprobados pertenecen a este grupo.

1. Características de un profármaco ideal

Un profármaco ideal será aquel donde el compuesto activo (el fármaco) se une a un transportador mediante un enlace covalente lábil (*carrier-linked prodrugs*) (Figura 7.1). Dicho enlace debe

ser lo suficientemente lábil, y al mismo tiempo lo suficientemente estable, para permitir la liberación efectiva del fármaco *in vivo.* Tanto el profármaco como el transportador no deben ser tóxicos una vez liberado el fármaco, ni conducir a metabolitos tóxicos. Por último, debe ser inactivo *in vitro.*

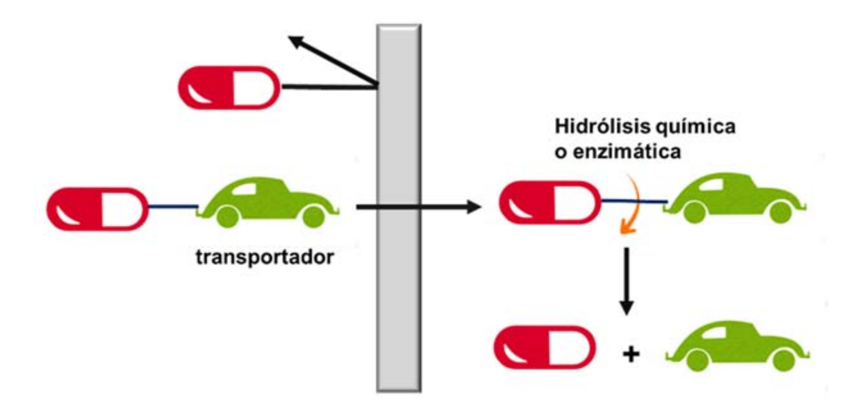


Figura 7.1: Representación del concepto de profármaco para superar o minimizar las barreras en el desarrollo de un fármaco.

El diseño de **profármacos** ofrece una alternativa válida para incorporar a la terapéutica aquellos compuestos que por alguna causa no pudieron pasar las fases clínica I, II o III; o que, habiéndolas pasado, en el posterior uso durante la fase IV, se observan efectos adversos o limitaciones. Es decir, los profármacos pueden rescatar a estructuras activas en los modelos biológicos pero que presentaron inconvenientes en el uso en humanos.

2. Clasificación de profármacos

Podemos dividir a los profármacos en dos clases: aquellos que se unen a un transportador y los que se denominan bioprecursores.

2.1 Profármacos unidos a un transportador

Resultan de la unión temporal de una molécula activa a un transportador (generalmente de naturaleza lipofílica) (Figura 7.2). Estos deben cumplir los requerimientos del profármaco ideal descritos anteriormente.

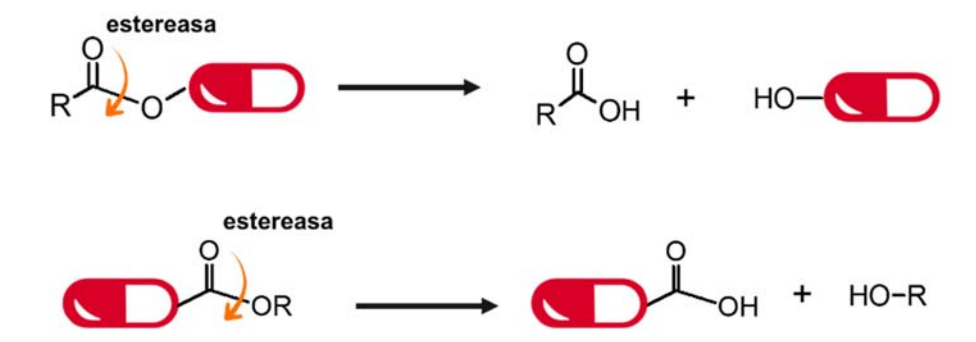


Figura 7.2: Concepto esquemático de la bioconversión de profármacos basados en un éster del ácido carboxílico correspondiente (reacción superior) y un éster de un alcohol (reacción inferior), cada uno de los cuales puede utilizarse como un transportador o fármaco.

2.2 Profármacos bioprecursores

Son aquellos que resultan de la modificación de la estructura molecular del fármaco y necesitan una activación metabólica en el organismo (ej. oxidación, reducción) para regenerar el principio activo.

Un ejemplo de este tipo de profármaco es el Omeprazol, un potente inhibidor de la secreción de ácido del estómago (Esquema 7.1), el cual es activado de forma selectiva en el medio ácido de la mucosa del estómago dando lugar a una sulfenamida tras una serie de reacciones químicas. Esta sulfenamida es la especie activa que presenta el efecto inhibidor en la bomba de protones.

Omeprazol
(profármaco)

sulfenamida
(especie activa)

ácido sulfénico

Esquema 7.1: Activación del omeprazol en el medio ácido del estómago.

3. Estrategias utilizadas para mejorar la biodisponibilidad de un fármaco

Con el fin de aumentar la biodisponibilidad de un fármaco existen diversas estrategias, entre las que podemos mencionar: (a) mejorar su solubilidad en agua, (b) mejorar su absorción intestinal pasiva o mediada por transportadores o (c) reducir o modular su metabolismo. Para lograr estos objetivos suelen adicionarse, eliminarse o enmascararse grupos funcionales capaces de presentar problemas.

Los cambios pueden conducir a un profármaco que luego libere el compuesto activo, o bien a una nueva entidad química activa con mejores características que el compuesto de partida. Por ejemplo, la polaridad se puede variar mediante la adición, eliminación o variación de sustituyentes hidrófobos adecuados. Como detallaremos más adelante, es posible reducir la polaridad de las moléculas al enmascarar un grupo funcional polar con un grupo alquilo o acilo. La polaridad disminuye no solo al modificar el grupo polar, sino también mediante la adición de un grupo alquilo grande que presenta un mayor efecto hidrofóbico. La adición de un sustituyente halógeno también aumenta la hidrofobicidad. En estos casos los sustituyentes más comúnmente utilizados son el cloro y el flúor, y con menor frecuencia el bromo. Es importante mencionar que se debe tener cuidado al enmascarar los grupos polares, ya que pueden ser útiles para la interacción del fármaco con su blanco molecular. Si este es el caso, resulta más beneficioso enmascarar temporalmente el grupo polar de manera que la máscara se elimine una vez que se absorba el medicamento, es decir diseñar un profármaco.

Por el contrario, si la molécula no es suficientemente polar pueden reemplazarse grupos alquilo grandes por más pequeños o se los elimina por completo. También es posible en algunos casos aumentar el tamaño de un grupo alquilo y disminuir el tamaño de otro.

Por ejemplo, el agente antifúngico Tioconazol sólo se usa para infecciones de la piel porque no es polar, presenta baja solubilidad acuosa y, por lo tanto, es poco soluble en sangre. La introducción de un grupo hidroxilo polar y más anillos heterocíclicos polares condujeron al agente antifúngico activo por vía oral Fluconazol, el cual presenta una mejor solubilidad acuosa y por lo tanto es útil en el tratamiento de determinadas infecciones sistémicas (Figura 7.3).

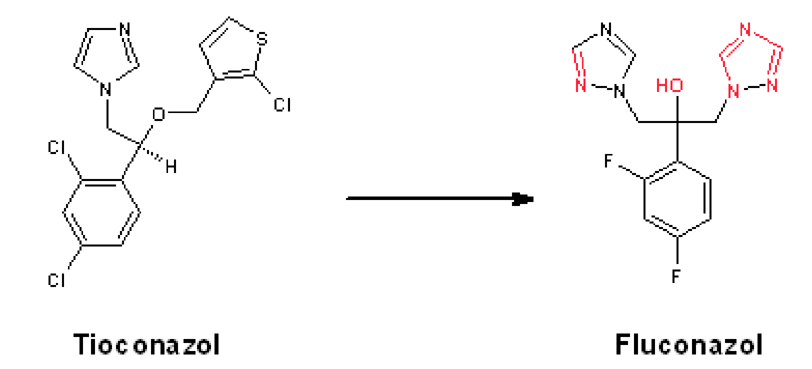


Figura 7.3: Modificación estructural del antimicótico de uso local Tioconazol que originó al Fluconazol, agente antifúngico utilizado en forma sistémica.

A continuación, describiremos las estrategias para mejorar la biodisponibilidad de un fármaco en el contexto de las prodrogas.

3.1 Mejora de la solubilidad en agua de un fármaco

La estrategia general para mejorar las características de solubilidad acuosa de un fármaco consiste en unir a este tipo de fármacos poco solubles, grupos moduladores que pueden ser cargados (ionizables) o macromoléculas neutras como el polietilenglicol (PEG). En el primer caso tenemos los ésteres de fosfatos, hemi-succinato y los ésteres de aminoácidos, entre otros.

El antiinflamatorio Dexametasona presenta una solubilidad en agua muy baja; el éster fosfato derivado presenta una mejor solubilidad en agua y se hidroliza rápidamente (presenta una vida media de tan solo diez minutos) liberando la especie activa *in vivo* (Figura 7.4.A).

El éster fosfato de Oxifenbutazona (Figura 7.4.B), utilizado como agente antiinflamatorio proporciona un nivel más alto en sangre del fármaco tanto en la administración oral como intramuscular, comparado con la administración de Oxifenbutazona a la misma dosis.

Figura 7.4: Profármacos de antiinflamatorios **A.** Éster fosfato de Dexametasona. **B.** Éster fosfato de Oxifenbutazona.

En el Esquema 7.2, la Prednisolona, un antiinflamatorio poco soluble en agua, aumenta su solubilidad cuando se forma el hemi-succinato (profármaco) originando la sustancia activa por hidrólisis.

Esquema 7.2: Profármaco hidrosoluble de Prednisolona.

La unión de aminoácidos a fármacos insolubles en agua por medio de un enlace éster es otra estrategia comúnmente empleada. A pH fisiológico, el aminoácido presente en el profármaco se encuentra ionizado, aumentando la solubilidad en agua de la molécula. Este tipo de estrategia presenta la ventaja de que, una vez activado el fármaco *in vivo*, el transportador liberado (un aminoácido o un oligopéptido) no presenta toxicidad en el organismo.

En cuanto a los grupos moduladores neutros para aumentar la solubilidad del fármaco en agua, usualmente se utilizan dioles, polioles y polímeros como los polietilenglicoles (PEG). Un ejemplo en la mejora de la solubilidad acuosa con grupos neutros se observa en los profármacos de Paclitaxel, un anticancerígeno, esterificado con PEG (Figura 7.5).

Figura 7.5: Fármaco Paclitaxel y su profármaco PEG-paclitaxel, el cual presenta una mejor solubilidad acuosa.

3.2 Mejora de la absorción intestinal pasiva de un fármaco

Para que un fármaco administrado por vía oral llegue a circulación sistémica tiene que atravesar la membrana del tracto gastrointestinal. La capacidad de un fármaco para atravesar dicha membrana está directamente relacionada con la lipofilia de la molécula. Por ello, aumentar la lipofilia de un fármaco es una estrategia empleada con frecuencia para mejorar el transporte pasivo en el intestino en casos de fármacos portadores de grupos polares como ácidos carboxílicos, fosfonatos, o de derivados nucleosídicos. En estos casos se diseñan profármacos que enmascaran los grupos polares ionizables, teniendo en cuenta el balance que debe haber entre la solubilidad en agua necesaria para su disolución en el tracto gastrointestinal y el aumento de lipofilia necesario para su posterior absorción por difusión pasiva.

Existen numerosos casos de profármacos diseñados mediante esta estrategia. Por ejemplo, los análogos de nucleósidos utilizados para el tratamiento de infecciones virales, los cuales presentan baja absorción por vía oral. Uno de ellos es el Aciclovir, para el cual se desarrolló el profármaco Valaciclovir (Figura 7.6). El cambio estructural consistió en generar una función éster

con valina, incrementando significativamente su solubilidad en agua, y aumentando su biodisponibilidad oral hasta cinco veces más, respecto al Aciclovir.

Con el mismo propósito se desarrolló el profármaco de la Zidovudina (AZT), utilizando el mismo aminoácido, obteniéndose el éster L-Val-AZT (Figura 7.6).

Aciclovir R = H

Valaciclovir: R= L-valina

AZT R=H

L-val-AZT R= L-valina

Figura 7.6: Profármacos de antivirales esterificados con aminoácidos para aumentar su absorción intestinal.

En el caso del antiviral Penciclovir, este presenta una biodisponibilidad del 4% cuando es administrado por vía oral en humanos. Sin embargo, su profármaco Famciclovir (Esquema 7.3), en el cual los dos grupos hidroxilo primarios se acetilaron y el grupo carbonilo de la posición 6 de la base fue eliminado, muestra un gran aumento en su biodisponibilidad oral (75%).

in vivo

Famciclovir

Penciclovir

Esquema 7.3: Transformación *in vivo* de Penciclovir a partir de su profármaco.

Otro ejemplo es el inhibidor enzimático Enalaprilato, un antihipertensivo cuya absorción fue mejorada mediante su conversión al profármaco Enalapril. En su forma activa se absorbe menos del 12%, mientras que el derivado inactivo (profármaco) presenta una absorción del 50-75%. La forma activa se libera *in vivo* mediante hidrólisis en el hígado. (Esquema 7.4).

Esquema 7.4: Hidrólisis enzimática del profármaco Enalapril a su forma activa Enalaprilato.

3.3 Modulación del metabolismo de un fármaco

La rápida descomposición metabólica del fármaco también puede protegerse mediante una estructura de profármaco. Esto se realiza mediante el enmascaramiento de los grupos funcionales metabólicamente lábiles. Este es el caso del Bambuterol, profármaco del broncodilatador Terbutalina, que presenta sus grupos fenólicos enmascarados en forma de carbamatos. Con la administración de una única dosis diaria del profármaco se consigue el mismo efecto que con la administración de tres dosis de Terbutalina. El profármaco es hidrolizado a Terbutalina por acción de colinesterasas no específicas (Esquema 7.5).

Esquema 7.5: Bioactivación de la prodroga de Terbutalina

Otro enfoque para mantener un nivel sostenido de un fármaco en sangre durante períodos de tiempo prolongados es asociar deliberadamente un grupo muy lipofílico al fármaco, generando así un profármaco. De esta manera, el profármaco tiende a almacenase en el tejido graso desde donde se libera de manera constante y lenta hacia el torrente sanguíneo, donde finalmente se convierte en el compuesto activo. Ejemplo de ello es el éster lipofílico del fármaco antipsicótico Flupenazina (Figura 7.7). El profármaco se administra por vía intramuscular y pasa lentamente desde el tejido graso a circulación sistémica, donde se hidroliza rápidamente generando la forma activa.

Figura 7.7: Profármaco de la Flupenazina

4. Profármacos que enmascaran la toxicidad del fármaco y efectos secundarios

Es posible usar profármacos para enmascarar los efectos secundarios y la toxicidad de los medicamentos. Por ejemplo, el ácido salicílico es un buen analgésico, pero causa sangrado gástrico debido al grupo fenólico libre. Esto se supera al enmascarar el fenol como un éster, generando la Aspirina, la cual se hidroliza posteriormente para liberar el fármaco activo (Esquema 7.6).

Esquema 7.6: Hidrólisis de la Aspirina.

Por otra parte, una de las limitaciones de los fármacos es la inespecificidad de su acción o falta de selectividad. Mediante el uso de profármacos es posible direccionar el compuesto activo hacia tejidos específicos, tal que el mismo se libere preferentemente en el sitio de acción, evitando así altas concentraciones sistémicas del compuesto activo, y de esa manera, reduciendo los efectos secundarios por interacciones inespecíficas. Por ejemplo, dirigir un fármaco al sistema nervioso central (SNC) es una tarea compleja, debido a la presencia de la barrera hematoencefálica, encargada, entre otras cosas, de impedir el paso de agentes externos desde el sistema sanguíneo al cerebro. Sin embargo, el conocimiento de los mecanismos de transporte y de la actividad enzimática en dicha barrera puede permitir elevar los niveles de un fármaco en el SNC empleando una aproximación profármaco adecuada. Esta estrategia se ha utilizado con la Dopamina, fármaco empleado en pacientes con la enfermedad de Parkinson, que no atraviesa la barrera hematoencefálica. El profármaco de Dopamina, denominado Levodopa, atraviesa

tanto el tracto gastrointestinal como la barrera hematoencefálica debido a que es sustrato del transportador de aminoácidos neutros LAT1 expresado en ambos sitios. Una vez en el tejido cerebral, la Levodopa es descarboxilada a Dopamina (Esquema 7.7). Adicionalmente, para evitar los efectos sistémicos de la Dopamina y aumentar su biodisponibilidad en el SNC (dado que sólo una pequeña parte de la Levodopa entra al cerebro), se administra asociada a inhibidores periféricos de la descarboxilasa como Beserazida y Carbidopa, aumentando su tolerancia y efectividad clínica.

Levodopa — dopa descarboxilasa, $-CO_2$ → Dopamina

Esquema 7.7: Conversión de Levodopa a Dopamina en el tejido cerebral.

Por otra parte, en la actualidad, uno de los retos en la quimioterapia del cáncer es el transporte selectivo del fármaco a las células tumorales sin afectar a las células sanas, como una forma de reducir los graves efectos adversos de los fármacos quimioterápicos. Las células tumorales presentan ciertas diferencias respecto a las células sanas, ya que su alta tasa de proliferación y su actividad biorreductora hacen que ciertas enzimas estén sobreexpresadas y puedan ser empleadas para la activación selectiva de profármacos en dichas células tumorales. Un buen ejemplo de profármaco activado por enzimas específicas de tumores es la Capecitabina, que se administra por vía oral a pacientes con cáncer de mama y colorrectal metastásico. La Capecitabina es un profármaco del 5-fluorouracilo que tras su absorción oral experimenta tres pasos de activación, dando lugar a altas concentraciones de 5-fluorouracilo en el tumor. Estas etapas de activación son: (a) hidrólisis en el hígado por acción de carboxilesterasas, (b) desaminación en el hígado y en células tumorales mediada por la citidina desaminasa y (c) liberación específica de 5-fluorouracilo en las células tumorales por acción de la enzima timidina fosforilasa (sobreexpresada en dichas células) (Esquema 7.8).

carboxiesterasa hepática

citidina desaminasa

hígado y tumor

timidina fosforilasa

tumor

capecitadina

5-fluorouracilo

Esquema 7.8: Bioconversión de Capecitabina a 5-fluorouracilo.

Referencias

Cabrera, S. (2010). Profármacos: pasado, presente y futuro. *Anales de Química.* 2010, 106(3), 207–214.

Patrick,G.L (2013). An Introduction to Medicinal Chemistry 5th Edition. Ed. Oxford University Press. ISBN: 978-0-19-969739-7.

Samaja,G. (2011). Noveles heterociclos con actividad anticonvulsivante derivados del Myo-inositol y Ácido Valproico (tesis doctoral). Recuperada del repositorio institucional de la Universidad Nacional de La Plata.

CAPÍTULO 8
Evaluaciones preclínicas en el descubrimiento de fármacos

Andrea V. Enrique

8.1 Ensayos preclínicos

Los estudios preclínicos constituyen un programa para evaluar la eficacia y seguridad de compuestos seleccionados con probabilidad de ser un fármaco. Este programa está sujeto a continuos cambios basados en el progreso del conocimiento científico, y se adapta a múltiples intereses de la comunidad científica, de la industria farmacéutica, y de las autoridades regulatorias.

Una vez identificados y seleccionados los compuestos candidatos a nuevos fármacos, se llevarán a cabo las pruebas experimentales para determinar los perfiles farmacodinámicos, farmacocinéticos y toxicológicos siguiendo los siguientes pasos:

8.1.1 Vehiculización de los fármacos

Es de suma importancia tener conocimiento de las propiedades fisicoquímicas del compuesto a prueba, especialmente determinar su solubilidad y estabilidad en el vehículo de administración. La insolubilidad de algunos compuestos impide alcanzar concentraciones efectivas cuando son administrados tanto por vías enterales como parenterales. En estos casos se deberá preparar soluciones o suspensiones en vehículos orgánicos inertes (aceites vegetales, polisorbatos, alcoholes, etc.) que pueden facilitar la administración de una dosis suficiente para ejercer el efecto deseado. Sin embargo, el empleo de estos disolventes puede ocasionar problemas como el propio efecto del vehículo, la precipitación del compuesto activo en el sitio de absorción o interacciones metabólicas entre el vehículo y la sustancia activa. Por ejemplo, un compuesto poco soluble en un medio lipídico puede cambiar de forma dramática su biodisponibilidad, según el solvente utilizado.

La estabilidad química del compuesto en solución o suspensión también debe ser conocida, debido a que los productos de degradación pueden interferir con los ensayos farmacológicos o ser confundidos con los metabolitos. Los métodos analíticos específicos, altamente sensibles y validados, son de gran utilidad para identificar químicamente al compuesto y a sus metabolitos, en fluidos y tejidos orgánicos.

8.2 Estudios farmacocinéticos

La farmacocinética se define como el estudio cuali y cuantitativo de la absorción, distribución, metabolismo y eliminación de un fármaco (ADME) en un organismo intacto.

El perfil farmacocinético que se pretende de un fármaco involucra una alta absorción, distribución adecuada, bajo metabolismo, eliminación completa del cuerpo y un riesgo toxicológico mínimo.

Durante la investigación preclínica es conveniente iniciar el estudio farmacocinético en un pequeño número de animales, al menos en dos especies, para determinar si el fármaco se absorbe por las vías de administración elegidas, conocer el rango de dosificación y la frecuencia de administración para obtener el efecto deseado. Aunque inicialmente no es necesario conocer el escenario completo de la biotransformación para decidir un estudio farmacodinámico, sí es conveniente tener una idea general de que enzimas participan en ella. La información farmacocinética es útil para predecir el inicio y la duración del efecto al establecer la velocidad y magnitud de la absorción, el destino del fármaco y/o de sus metabolitos en los diferentes órganos, los procesos de transformación del fármaco, la velocidad y vías de eliminación.

Los estudios ADME son llevados a cabo mediante modelos *in silico*, *in vitro* e *in vivo.* Los modelos *in vitro* proveen parámetros ADME como permeabilidad aparente, estabilidad metabólica, reacción fenotípica, unión a proteínas, partición sangre-plasma, potencial interacción fármaco-fármaco, citotoxicidad y proliferación celular, e inhibición hERG.

Los modelos *in vivo* permiten conocer la biodisponibilidad oral, distribución, aclaramiento, y duración de exposición al fármaco y sus metabolitos. Finalmente, los estudios *in silico* predicen el comportamiento del fármaco basándose en sus propiedades fisicoquímicas en combinación con simulación de la interacción con la estructura 3D de una proteína (una enzima o un transportador).

Hay una gran variedad de modelos que permiten evaluar las propiedades ADME de un fármaco, algunos de ellos se encuentran listados en la Tabla 8.1. Durante las etapas tempranas y medias del descubrimiento de fármacos (identificación de leads y optimización) los estudios ADME son conducidos generalmente en modo de screening de alta performance.

Tabla 8.1: Modelos utilizados en estudio ADME típicos.

Estudios ADME	Sistema de estudio
Estabilidad metabólica	Preparado de hígado, enzimas
Identificación de metabolitos	Preparado de hígado, birreactor, ensayos *in vivo*
Reacción fenotípica	Microsomas, hepatocitos, enzimas
Inhibición de CYP	Microsomas, hepatocitos, enzimas
Innducción de CYP	Microsoma, *ex vivo*
Tansportadores	Lineas celulares Ej: CACO-2
Unión a proteínas plasmáticas	Plasma
Balance de masa	Animales y humanos
Perfil metabólico	Animales y humanos
Disposición	Sujetos sanos o pacientes
Comparación de especies	Animales y humanos
Distribución en tejidos	Ratas

Entre las técnicas de screening de alto rendimiento usadas para recopilar datos farmacocinéticos de múltiples candidatos a fármacos en un solo experimento se encuentra el Cassette dosing, o dosificación en casete. El típico estudio farmacocinético cassette dosing involucra la administración simultánea de 5 a 10 compuestos a cada animal de un grupo de animales. Se obtienen muestras de sangre en serie y se utilizan técnicas de cromatografía de líquidos con espectrómetro de masas para medir la concentración de los analitos en cada muestra. Por lo tanto, en lugar de ejecutar 5-10 estudios de farmacocinética separados, se puede obtener una cantidad similar de información de un solo estudio. Esta metodología puede facilitar la evaluación rápida de muchos compuestos con menor número de animales, aunque también presenta inconvenientes:

En primer lugar, generalmente los 5-10 candidatos evaluados son de estructura química similar. Debido a las similitudes estructurales, las interacciones metabólicas de fármacos son más probables. Es probable que las estructuras comunes usen vías metabólicas similares que pueden conducir a la saturación de las mismas, lo que da lugar a bajos valores de aclaramiento. Esto puede sugerir que los fármacos tienen tiempos de residencia más largos de lo que se esperaría cuando se dosifican por separado.

En segundo lugar, la dosificación oral presenta el problema adicional de las interacciones relacionadas con la absorción y el metabolismo de primer paso, por lo que es preferible usar la

dosificación intravenosa (IV) para el experimento. Sin embargo, en muchos casos la solubilidad del compuesto impide la preparación de una formulación IV aceptable.

En tercer lugar, los desafíos analíticos se magnifican con la dosificación en casete. Los avances en la espectrometría de masas en los últimos 15 años han reducido la necesidad de separación cromatográfica en una columna. Sin embargo, con compuestos estructuralmente similares que difieren por pequeñas diferencias de peso molecular, puede haber problemas con la identificación de cada uno de ellos. Estas cuestiones pueden resolverse, pero por lo general requieren un trabajo analítico adicional para asegurar que se produzca una separación temporal de los analitos en la columna.

La dosificación de casete ha generado controversias debido a las preocupaciones respecto a los errores que posee. Algunos informes afirman que se obtienen datos farmacocinéticos fiables, pero sólo unos pocos han demostrado realmente una correspondencia razonable de los parámetros farmacocinéticos obtenidos de la dosificación de casete y la dosificación convencional de un solo compuesto. Sin embargo, esta técnica sigue siendo una herramienta útil para evaluar la farmacocinética comparativa de una serie de compuestos. Ha sido utilizada satisfactoriamente para evaluar la farmacocinética de un amplio rango de estructuras químicas en múltiples especies incluidas ratas, ratones, perros y monos.

8.3 Estudios farmacodinámicos

Durante los estudios preclínicos, la evaluación farmacodinámica de los compuestos tiene como objetivo: a) demostrar en modelos experimentales *in vitro* e *in vivo* el efecto terapéutico previsto del nuevo fármaco, y b) establecer la relación entre la dosis administrada con la intensidad y duración de los efectos producidos.

La relación entre la intensidad de un efecto particular y los logaritmos de las dosis administradas representadas en un plano cartesiano grafican una curva dosis-respuesta, la cual adquiere una forma sigmoidal. Es de particular interés examinar la forma de la curva, las respuestas a bajas dosis y altas dosis del fármaco y la pendiente de la porción recta de la curva.

Los estudios farmacodinámicos deberán hacerse extensivos a los principales órganos y sistemas de los animales para determinar, no solo el efecto terapéutico previsto, sino también revelar otros posibles efectos terapéuticos o nocivos sobre los sistemas cardiovascular, respiratorio, nerviosos central, nerviosos vegetativo, neuromuscular, urinario, endócrino, digestivo, entre otros sistemas de importancia farmacológica. Además, deberá describirse el o los mecanismos de acción responsables de los efectos encontrados durante el estudio.

El tipo y la cantidad de los ensayos farmacodinámicos iniciales dependerán del uso que se pretende dar al nuevo fármaco. Por ejemplo, si se pretende seleccionar fármacos para el tratamiento de enfermedades infecciosas bacterianas se evaluará, en primera medida, la actividad contra organismos infecciosos, o bien, si se pretende descubrir fármacos para enfermedades tromboembólicas, los ensayos preclínicos seleccionados evaluarán la capacidad de, por ejemplo, prolongar el tiempo de coagulación.

8.3.1 Perfil farmacológico

En cada modelo experimental al que el compuesto se somete a prueba, se observa un comportamiento distintivo que recibe el nombre de "perfil farmacológico". El compuesto a prueba presenta tantos perfiles farmacológicos como estudios realizados, y el análisis del conjunto de estos perfiles determina si cumple con los criterios específicos necesarios para proseguir la investigación del fármaco.

Tomemos como ejemplo un fármaco que se espera actúe como antagonista de receptores H_1. Se realizarían los siguientes estudios:

- A nivel molecular se estudiará la selectividad y afinidad del compuesto en receptores clonados, en membranas de líneas celulares cultivadas y en tejidos que expresen el receptor H_1.
- Actividad enzimática, por ej. Citrocromo P450 hepático, para saber si es sustrato y si tiene efectos inductores o inhibidores de estas enzimas.
- Efectos sobre el funcionamiento celular relacionados con la estimulación del receptor H_1, para determinar que el fármaco en estudio es un antagonista H_1 y, es eficaz para inhibir las funciones mediadas por dichos receptores.
- Se realizarán pruebas *in vivo* sobre la eficacia y duración del efecto después de administrar el compuesto en comparación con fármacos de referencia por ej: Clorfeniramina, Difenhidramina, Terfenadina, Aztemizol. Se mediría por ejemplo a la inhibición de mediadores inflamatorios (histamina, prostaglandinas, leucotrienos, bradicininas entre otros)
- Comparación con fármacos de referencia en estudios *in vitro* en otros tejidos como músculo liso gastointestinal y bronquial.

En cada paso, el compuesto tendrá que cumplir con criterios de "comportamiento" específicos para poder proseguir con los estudios.

Tal como en el ejemplo mencionado, la batería de ensayos que se realiza a un nuevo fármaco puede incluir estudios *in vitro* y estudios *in vivo.*

Estudios *in vitro*

Los estudios *in vitro* (del latín: *dentro del vidrio*) se refieren a técnicas para realizar experimentos en un ambiente controlado fuera de un organismo vivo. Apuntan a describir los efectos de una variable experimental en un subconjunto de las partes constitutivas de un organismo, con una tendencia a enfocarse en órganos, tejidos, células, componentes celulares, proteínas y/o biomoléculas. Son apropiados, por ejemplo, para deducir un mecanismo de acción, debido a que presenta menos variables y reacciones amplificadas, obteniendo resultados más discernibles. La principal ventaja de los modelos in vitro es que permiten un enorme nivel de simplificación del sistema de estudio, por lo que el investigador puede centrarse en un pequeño número de componentes. Por ejemplo, la identidad de las proteínas del sistema inmune (por ejemplo,

anticuerpos), y el mecanismo por el cual se reconocen y se unen a los antígenos extraños, seguiría siendo muy oscuro si no fuera por el uso extensivo de técnicas *in vitro* para aislar las proteínas, identificar las células y los genes que las producen, el estudio de su interacción con los antígenos, e identificación de cómo las interacciones conducen a las señales celulares que activan otros componentes del sistema inmune.

Otra ventaja de los ensayos *in vitro* es que requieren pequeñas cantidades de los compuestos a ensayar.

La principal desventaja de los estudios experimentales *in vitro* es que a veces puede ser muy difícil extrapolar los resultados a la biología del organismo intacto. Los investigadores deben ser cuidadosos para evitar la mala interpretación de sus resultados, que a veces puede llevar a conclusiones erróneas acerca de organismos y biología de sistemas.

Estudios *in vivo*

In vivo (latín: *dentro de lo vivo*) significa “que ocurre o tiene lugar dentro de un organismo”. En ciencia se refiere a experimentación hecha dentro de un organismo vivo completo, usualmente animales, el ser humano o plantas.

La capacidad de reproducir enfermedades humanas en modelos animales presenta una gran ventaja para la medicina experimental moderna. El modelo animal ideal es un modelo que reproduce el desorden humano en todos los aspectos. Alternativamente, el modelo animal podría ser isomorfo, cuando el trastorno se replica pero no la etiología subyacente (que en muchas enfermedades es insuficientemente conocida o no se conoce en absoluto); o predictivo, cuando no se asemeja al trastorno humano, pero permite predicciones acerca del mismo o de la respuesta al tratamiento.

8.4 Estudios de toxicidad

Los compuestos que pasan satisfactoriamente por los procedimientos iniciales de selección y perfil farmacológico deben evaluarse para identificar los efectos nocivos que producen sobre las estructuras y funciones de los sistemas vivos. El objetivo de los estudios de toxicidad es establecer dichos efectos en diferentes especies animales, anticipar posibles efectos tóxicos en el hombre y decidir si el nuevo fármaco es razonablemente seguro para su experimentación clínica.

Aunque ningún producto químico es completamente seguro (debido a que todas las sustancias son tóxicas y solo la dosis determina la toxicidad), con frecuencia es posible estimar el riesgo o peligro potencial que un agente químico puede ocasionar sobre la salud humana cuando es objeto de exposición aguda o crónica. Para este propósito se utilizan animales expuestos a dosis únicas (toxicidad aguda) o repetidas (subaguda, subcrónica y crónica) del fármaco por la vía de administración que se pretende emplear en el humano. Si se piensa administrar el fármaco por varias vías se deben hacer los estudios correspondientes en cada vía empleada, al menos de toxicidad aguda y subaguda (hasta 4 semanas de exposición). En los

estudios a largo plazo de toxicidad crónica y carcinogénesis, el fármaco se administra de la forma menos traumática posible, por ejemplo, junto con el alimento, si la administración es oral, o mediante una cánula, si es parenteral.

Para la elección de las dosis tóxicas se lleva a cabo un estudio preliminar agudo, en un número limitado de animales (Ej. 10 ratas de cada sexo). Las agencias regulatorias exigen que se establezca la menor dosis en la que aparecen efectos adversos relacionados con la administración del fármaco y una dosis tóxica que ocasione efectos graves, sin producir más que una reducida mortalidad de los animales. Los animales sobrevivientes serán sometidos a exámenes anatomopatológicos y bioquímicos. Los animales tratados con el fármaco problema se distribuirán al menos en 5 grupos que recibirán dosis relativamente crecientes.

En todos los estudios con dosis repetidas el número de animales es importante. Las agencias regulatorias exigen que al final del período de tratamiento (que pueden ser de hasta 2 años para estudios de carcinogénesis) quede un número de animales sobrevivientes lo suficientemente amplio que permita comparaciones estadísticas entre grupos. Debe tenerse en cuenta que la muerte de los animales puede deberse no solo a un efecto tóxico del fármaco, sino también a accidentes operativos o incluso a espontáneamente, teniendo en cuenta que la esperanza de vida de un ratón es de alrededor de 2 años.

En cuanto a la especie a utilizar, las agencias regulatorias recomiendan que se utilicen aquellas especies cuyo metabolismo se asemeje más al del humano. En general, los estudios de toxicidad aguda se llevan a cabo en roedores (rata y ratón), los estudios de toxicidad subaguda, subcrónica y crónica en rata y perro (en algunos casos el perro es sustituido por el mono), y los estudios de carcinogénesis en rata y ratón. Los estudios de fertilidad y toxicidad peri- y postnatal se suelen hacer en rata y la teratogenia en rata y conejo. Los estudios de mutagénesis se realizan en sistemas celulares *in vitro* (bacterias, células de mamífero o linfoma) o *in vivo* en el ratón (prueba del micronúcleo).

8.4.1 Toxicidad aguda

El objetivo es evaluar los efectos de una elevada dosis del fármaco. Usualmente le punto final del estudio es la muerte del animal. La toxicidad aguda se expresa en Dosis Letal 50 (DL_{50}), que representa la dosis del fármaco que provoca la muerte del 50% de los animales evaluados y permite establecer la potencia de toxicidad del fármaco. La observación de los animales se lleva a cabo después de la administración de la sustancia y dura hasta 14 días, después de los cuales, los animales son sacrificados y autopsiados para observar efectos locales.

La determinación de la DL_{50} suele llevarse a cabo en ratones en, al menos, dos vías de administración. El tamaño pequeño del ratón permite utilizar dosis modestas del nuevo fármaco, y, a fin de satisfacer los requisitos de las autoridades regulatorias, es frecuente que se realicen estudios de toxicidad aguda en otra especie roedora (generalmente ratas) y al menos en una especie no roedora (generalmente perros). Esto se realiza a fin de obtener evidencia preliminar sobre diferencias interespecies en la respuesta al fármaco. Cuando se utilizan especies no roedores el punto final del experimento no es la muerte del animal sino la aparición de efectos adversos graves. La Tabla 8.2 muestra el grado de toxicidad de un fármaco en función del valor de DL_{50}.

Tabla 8.2: Potencia de la toxicidad de un fármaco en función de la DL50

Potencia	DL_{50} (mg/kg)
Extremadamente tóxico	< 1
Altamente tóxico	5-50
Moderadamente tóxico	50-500
Ligeramente tóxico	500-5000
Prácticamente no tóxico	5000-15000

8.4.2 Toxicidad subaguda

Las autoridades sanitarias exigen que, antes de autorizar la administración de una dosis única al ser humano, la realización de estudios de toxicidad subaguda en dos especies animales, una de las cuales debe ser no roedor. En esta prueba, la administración del fármaco se lleva a cabo diariamente mediante período que oscilan entre 15 días y 4 semanas. Se suelen utilizar entre 4 y 5 dosis de sustancia (vehículo, una dosis baja, dos dosis intermedias y una dosis alta). En la rata se requieren al menos 10 animales de cada sexo para cada dosis y en el perro al menos 4 de cada sexo. Durante el estudio se controlan diariamente ciertos parámetros: aspecto, comportamiento, peso, consumo de agua y alimento, ECG, examen oftalmoscópico, etc., y al final los animales son sacrificados y autopsiados. Al inicio del estudio y antes de la autopsia, se toman muestras de sangre y orina para ser analizadas. La autopsia consiste en el examen macroscópico y anatomopatológico de vísceras y tejidos.

8.4.3 Toxicidad subcrónica y crónica

Las agencias regulatorias exigen la evaluación en al menos dos especies animales, una de las cuales deberá ser no roedor, para los estudios de toxicidad subcrónica y crónica. Los estudios de toxicidad subcrónica suelen durar tres meses mientras que los de toxicidad crónica suelen durar de 6 meses a 1 año, según el uso terapéutico de la sustancia. Si, por ejemplo, se trata de un antihipertensivo o antidiabético oral, cuya administración en el hombre se prevé durante años, la toxicidad crónica durará un año. En el caso de un antibiótico, en cambio, usado en el hombre durante períodos de 8-10 días, estudios de 6 meses serán suficientes.

8.4.4 Estudios de carcinogénesis

Los estudios de carcinogénesis determinan si la exposición prolongada al fármaco problema promueve la aparición de tumores por afección de factores de crecimiento o activación de oncogenes. Se exige que estas pruebas se realicen en dos especies animales, típicamente ratas y ratones con una duración que oscila entre 18 y 24 meses.

8.4.5 Estudios de mutagénesis

Se requiere de tres ensayos mutagénicos *in vitro* y uno *in vivo.* Para los ensayos *in vitro* se requieren células bacterianas (*Salmonella typhimorium*) o de mamífero (linfocitos humanos) que se incuban con concentraciones crecientes del fármaco a ensayar en presencia y ausencia de la fracción S-9. Esta fracción consiste en una suspensión de microsomas hepáticos de rata que se añaden para favorecer la posibilidad de que un fármaco, que en un principio no es mutagénico, se metabolice a un producto que si lo sea. En los ensayos *in vivo* se observan posibles defectos durante la hematopoyesis en ratón (prueba del micronúcleo).

8.4.6 Estudios sobre la reproducción

Los estudios de reproducción se realizan en ratas y conejos y tienen como objetivo determinar el efecto de los fármacos sobre el ciclo reproductivo.

En los estudios de fertilidad, las ratas machos y hembras, son expuestas al nuevo fármaco desde 60 y 15 días respectivamente antes del apareamiento hasta el destete de las crías.

En los estudios de teratogenia, la administración del fármaco tiene lugar desde el momento de la implantación del óvulo hasta el parto, esto es desde el día 8 al 21 de gestación.

Finalmente, los estudios de toxicidad peri y postnatal tienen por objeto determinar posibles efectos sobre el desarrollo de los embriones, el parto y la lactancia. En algunos casos determinados, estos estudios se extienden a varias generaciones. Todos los datos relativos a estos estudios, así como muestras de tejidos en formol y preparados microscópicos, deben conservarse un mínimo de 10 años.

8.5 Uso de animales de laboratorio

El animal de laboratorio es una de las piezas fundamentales en las ciencias biomédicas. Son usados como modelos para investigar y comprender las causas, diagnóstico y tratamiento de enfermedades que afectan al humano y a los animales, además de sus importantes aportes en la docencia biológica y en el desarrollo, producción y control de medicamentos, alimentos y otros insumos, donde en muchos casos hasta la fecha son insustituibles.

El uso de los animales de laboratorio en las investigaciones biomédicas representa un elemento fundamental en el desarrollo de importantes avances en la prevención y tratamiento de las enfermedades transmisibles y no transmisibles. Basta recordar las vacunas de la rabia, viruela, tétanos, difteria, tos convulsa y poliomielitis; el desarrollo de diversos antibióticos, la insulina y el conocimiento de las bases genéticas de la herencia. Los avances de la investigación en cáncer, cardiología, trasplantes de órganos, síndrome de inmunodeficiencia adquirida, enfermedad de Alzheimer, etc. se deben también a las contribuciones de los estudios realizados en animales de laboratorio. Estos animales son para el investigador un reactivo biológico, por lo que su pureza debe ser vigilada, controlada y contrastada, al igual que cualquier reactivo, sin descui-

dar su posible contaminación biótica. Por esta razón se requiere la producción de animales "estandarizados o definidos" con características genéticas y sanitarias definidas, criados en ambientes controlados que respeten los requerimientos de la especie, con el correcto cumplimiento de los principios éticos y de bienestar animal.

Dentro de los modelos de experimentación sin duda el ratón es el más conocido y utilizado en la mayor parte de las experiencias *in vivo*. Es en general el modelo elegido para conocer la reacción de un organismo mamífero frente a una agresión, una intoxicación o una infección experimental (parasitaria, bacteriana o vírica); o a reacciones o trastornos inmunológicos, oncología, teratología y embriología. Muchos investigadores consideran al ratón como un modelo animal casi perfecto porque además de su corto tiempo generacional, alta performance reproductiva y fácil mantenimiento, son los animales más sofisticados que pueden ser utilizados por los investigadores. La rata ocupa el segundo lugar, la cual se emplea además en investigaciones nutricionales, comportamentales y endocrinológicas. El conejo es utilizado fundamentalmente en la producción de antisueros; farmacología; toxicología; teratogenicidad y reproducción. Los cobayos son modelos de estudios inmunológicos, farmacológicos y nutricionales. El hamster se utiliza fundamentalmente en la reproducción, citogénesis e inmunología.

En la actualidad se ha desarrollado y perfeccionado una gran variedad de modelos animales, a fin de cubrir las exigencias de los investigadores para llevar a cabo los experimentos cada vez más sofisticados. Existen más de 478 cepas consanguíneas de ratón y 234 de rata de laboratorio. Dentro de éstas, las hay de uso general o especial, como los modelos para enfermedades autoimnunes, endócrinas y tumorales. También se destacan los híbridos F1, cepas congénitas, coisogénicas, transgénicos y *knock-out*.

Idealmente, el modelo debe ser homólogo, es decir, que los signos y la evolución de la condición que se está estudiando en el animal sean idénticos a los de la especie original, aunque esto no es muy frecuente y se suelen utilizar modelos parciales, no capaces de imitar completamente el proceso, pero permiten el estudio de ciertos aspectos de la condición evaluada.

8.5.1 Principios éticos

En 1959 William Russell y Rex Burch, en Inglaterra, en su famoso libro "The principles of humane animal experimental techniques", expusieron por primera vez que la excelencia científica está fuertemente ligada al uso humanitario de los animales de laboratorio. Definen claramente las normas en las que se basan los principios éticos en la investigación con animales: las tres ***"Rs": Reducir, Reemplazar y Refinar***. Estos son los fundamentos para una racional e inteligente estrategia que tiene como objetivo minimizar el uso de animales, las causas de dolor y de distrés. Los proyectos de investigación que requieren el uso de animales de laboratorio deben ser realizados con el número mínimo necesario de animales que permitan obtener resultados científicamente válidos. El perfeccionamiento del diseño de los experimentos y la selección del modelo más adecuado, contribuyen al cumplimiento de este principio. Los procedimientos in vivo deben ser reemplazados siempre que sea posible por métodos alternativos que no usen animales vivos, como modelos matemáticos, simulación por computadora, test serológicos, cultivos

celulares y sistemas biológicos in vitro. El refinamiento involucra fundamentalmente la estandarización según parámetros internacionales: definición genético-sanitaria y la calidad del ambiente donde son criados y mantenidos los animales antes y durante la experimentación. Incluye todos los procedimientos para minimizar y eliminar el dolor, así como todos los métodos de enriquecimiento ambiental para asegurar el bienestar animal. Estos principios han sido adoptados para su aplicación a través de normas bioéticas, tales como la Declaración de la Asociación Médica Mundial Sobre el Uso de Animales en la investigación Biomédica, adoptada por la 41ª Asamblea Médica Mundial celebrada en Hong Kong, en 1989; la Guía del Consejo Internacional de Organizaciones de Ciencias Médicas para la investigación biomédica que involucre animales; las Guías para el cuidado y uso de los animales de laboratorio de los Institutos Nacionales de Salud de Estados Unidos.

El diseño del experimento es solo una parte del desarrollo de la investigación, su éxito también depende de las condiciones de vida del animal y de la calidad de los recursos humanos (especialización y acreditación). Para el diseño del experimento se han especificado una serie de pautas y recomendaciones a seguir, como:

1. Seleccionar el modelo animal adecuado, el cual depende de la especie, cepa y de la calidad del animal.

2. Todas las investigaciones deberán justificar el número de animales seleccionados. El cálculo de la muestra se hará por procedimientos estadísticos y la misma no deberá superar el mínimo necesario para asegurar la confiabilidad de los resultados. El número de animales a utilizar depende de las características mencionadas anteriormente y del protocolo de experimentación. La selección del inóculo: dosis, vía y frecuencia de inoculación y la determinación del punto final del experimento, son elementos que también se deben definir.

3. Realizar los procedimientos experimentales adecuados. Estos están condicionados a los conocimientos y calificación de los recursos humanos. Involucran desde la experiencia a la profesionalidad en el método a aplicar, además del reconocimiento del dolor, al manejo del alivio y la sedación. Cualquier estudio con animales puede producir en estos distrés o dolor, por lo que cualquier esfuerzo es poco para tratar de minimizarlo sin afectar el objetivo del experimento. Hasta el momento, se considera que cualquier procedimiento que cause dolor o distrés en humanos, también puede causarlo en los animales. El incremento o disminución de la defecación y de la orina, la pérdida de peso, los cambios en el comportamiento y la inmovilidad son algunos de los signos clínicos. En roedores, son indicadores del dolor los cambios de actitud y de la apariencia del pelo corporal.

4. La realización de pruebas piloto es fundamental para madurar y definir la investigación, en lo que concierne a la muestra y sus procedimientos. La prueba piloto posibilita la elaboración del protocolo y de cuantos protocolos sean necesarios para el registro disciplinado de las observaciones y de los datos obtenidos durante las diversas etapas de la experimentación.

5. Especificar el método de eutanasia y definir el punto final humanitario del experimento. La eutanasia se define como la muerte sin dolor y distrés. La verdadera eutanasia supone una insensibilización rápida, mantenida hasta que se produzca la muerte. Se aplica al final del protocolo

o puede ser necesaria antes de finalizado el mismo, por razones de dolor o distrés, en todo caso que no pueda ser aliviado con analgésicos o sedantes u otros tratamientos.

La mayoría de los estudios preclínicos se realizan en las instalaciones de las grandes compañías farmacéuticas. Sin embargo, suelen recurrir a centros especializados de investigación por contrato (CROs) que cuentan con el material y la tecnología necesaria para la realización de estudios complejos.

En Argentina contamos con una plataforma tecnológica (PPL) financiada por el Ministerio de Ciencia, Técnica e Innovación Productiva, conformada por la Universidad de Buenos Aires, el Consejo Nacional de Investigaciones Científicas y Técnicas y la Universidad Nacional de La Plata (EBAL). Esta plataforma tiene como objetivo mejorar la calidad de los animales de laboratorio que se emplean en nuestro país y realizar ensayos preclínicos acorde con la normativa de la Administración Nacional de Alimentos, Medicamentos y Tecnología Médica (Disposición 6677/2010) y con los criterios fijados por la *International Conference for the Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use* (ICH) bajo el marco de las Buenas Prácticas de Laboratorio (BPL) (http://www.ebal.com.ar/index.php)

Los estudios preclínicos se llevan a cabo según las buenas prácticas de laboratorio (GLP), en condiciones ambientales muy especiales para garantizar que no interfieran factores ajenos al fármaco ensayado y con la participación multidisciplinaria de farmacólogos, veterinarios, farmacéuticos, químicos analíticos, anatomopatólogos, etc. Sin embargo, el uso de animales, particularmente en estudios de toxicología ha sido criticado debido a la percepción de crueldad que conllevan. En respuesta a este reclamo, y con el propósito de reducir el número de animales sacrificados en los estudios preclínicos, se recomienda a los grupos de investigación no duplicar estudios y evitar aquellos que son irrelevantes en el hombre por extrapolación. Además, en los casos en que sea posible se deben utilizar métodos alternativos, como ensayos en cultivos de diferentes líneas de células humanas, que incrementen la calidad de los ensayos y de la información. Por último, los objetivos actuales de la farmacología preclínica han sido redefinidos hacia propósitos más humanos como, por ejemplo, que la muerte no sea un objetivo en los estudios de toxicidad aguda.

8.5.2 Bienestar animal

El animal de laboratorio es considerado un “reactivo biológico”, es decir, un “animal de experiencia, en función del tema de estudio, capaz de dar una respuesta fiable y reproducible, cuya pureza debe ser vigilada, controlada y contrastada lo mismo que otro reactivo (químico o físico)”. Por lo tanto, es necesario escoger el modelo animal apropiado, emplear individuos homogéneos y evitar cualquier contaminación durante su utilización para reducir posibles errores en los resultados y asegurar la fiabilidad de la información recopilada. Sin embargo, el concepto de bienestar animal incorpora la necesidad de un “estado de salud mental y físico, donde el animal está en perfecta armonía con el ambiente que le rodea”. En la búsqueda del bienestar animal de plantea que los mismos deben gozar de **5 libertades**: libertad de hambre y sed; libertad de incomodidad

y dolor; libertad de lesión y enfermedad; libertad de miedo y angustia y libertad de expresar patrones de conducta esenciales para la especie. Tanto la última de las 5 libertades como el principio de Refinamiento tienen directa relación con el **enriquecimiento ambiental**, que se define como cualquier modificación en el ambiente de animales cautivos con el fin de proveerle estímulos y mejorar su bienestar físico y psicológico. Animales privados de realizar conductas naturales específicas para su especie han mostrado signos de sufrimiento como alteraciones conductuales, estrés crónico y otras condiciones patológicas que pueden perturbar los resultados obtenidos en una investigación. Existen enriquecimientos sociales, físicos, sensoriales, nutricionales, ocupacionales y cognitivos. El enriquecimiento a seleccionar debe ser específico para cada especie, considerando también características como raza, género, edad, estatus social, tipo de interacción del animal con su medio y variaciones individuales, por lo que los elementos y programas incorporados deberían estar comprobados y validados para tal fin.

Referencias

Mendoza Patiño N. (2008). Farmacología Preclínica en Farmacología Médica. Editorial Médica Panamericana S.A. México D.F.

Hernández, S. (2006). El modelo animal en las investigaciones biomédicas. *Biomedicina*, 2(3), 252-256.

Vogel, H. S.G. (2008). Drug Discovery and Evaluation: Pharmacological Assays. 3ra Ed. Editorial Springer. Alemania.

Los autores

Coordinadora

Gavernet, Luciana

Licenciada en Química (2001) y Doctora de la Facultad de Ciencias Exactas – Universidad Nacional de La Plata (UNLP) (2006). En 2009 ingresó a la Carrera del CONICET y actualmente reviste la categoría de Investigadora Adjunta. Desde 2018 es Profesora Titular a cargo de la asignatura Química Medicinal de la Facultad de Ciencias Exactas (UNLP). Ha publicado más de 30 artículos en publicaciones científicas y más de 70 presentaciones a congresos. Ha recibido subsidios de financiamiento científico y premios entre los cuales se encuentra el premio "Ciencia, Tecnología e Innovación 2017" otorgado a investigadores por la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires por el proyecto: Inhibidores de Anhidrasa Carbónica para el Tratamiento de la Epilepsia Febril. Ha dictado diversos cursos de postgrado y conferencias en ámbitos nacionales e internacionales. E-mail: lgavernet@biol.unlp.edu.ar

Autores

Alberca, Lucas Nicolás

Licenciado en Biotecnología y Biología Molecular egresado de la Facultad de Ciencias Exactas - UNLP (2014). En el año 2014 comenzó sus estudios de postgrado en la Facultad de Ciencias Exactas, UNLP y recibió una beca doctoral del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Es Ayudante Diplomado de la Cátedra de Química Medicinal (Facultad de Ciencias Exactas, UNLP). Ha recibido becas y subsidios de viaje de diversas instituciones internacionales. Sus contribuciones científicas incluyen 7 publicaciones en revistas científicas internacionales y 14 presentaciones a congresos nacionales e internacionales. Ha recibido 3 premios por sus trabajos científicos. E-mail: lalberca@biol.unlp.edu.ar

Bellera, Carolina Leticia

Farmacéutica y Lic. en Cs. Farmacéuticas egresada de la Facultad de Ciencias Exactas - UNLP (2007). En 2014 obtuvo el título de Doctora de la Facultad de Ciencias Exactas - UNLP. Es Pro-

fesora Adjunta de las Cátedras de Química Medicinal y Nuevas Estrategias Aplicadas al Descubrimiento de Fármacos - Facultad de Ciencias Exactas - UNLP. En 2016 ingresó a la Carrera del Investigador Científico del CONICET. Ha recibido subsidios de financiamiento científico y premios destacados entre los cuales se encuentra el Premio a la Innovación de UNLP en 2015. Sus contribuciones científicas incluyen unas 25 publicaciones nacionales e internacionales, incluyendo artículos originales, artículos de revisión, capítulos de libro y patentes de invención, mayormente vinculados al desarrollo de nuevos tratamientos para la enfermedad de Chagas y temáticas biofarmacéuticas. E-mail: cbellera@biol.unlp.edu.ar

Di ianni, Mauricio Emiliano

Farmacéutico egresado de la Facultad de Ciencias Exactas - UNLP, es Ayudante Diplomado en la Cátedra de Química Medicinal desde el año 2012 y Jefe de Trabajos Prácticos en Biofarmacia y Farmacotecnia I desde 2018. Realizó su doctorado en el actual Laboratorio de Investigación y Desarrollo de Bioactivos (LIDeB) haciendo uso de una beca doctoral del CONICET (2011-2015), obteniendo el título de Doctor de la Facultad de Ciencias Exactas en diciembre de 2015. Es autor de 12 trabajos científicos publicados y de 2 capítulos de libro. Premio Dr. Enrique herrero Ducloux 2016, a la mejor tesis en el área biotecnología, Asociación Química Argentina (AQA) y egresado distinguido de posgrado de la UNLP 2016. Actualmente es becario posdoctoral del CONICET e investiga en el desarrollo de nuevos nanovehículos farmacéuticos para el tratamiento de la epilepsia refractaria. E-mail: mdi@biol.unlp.edu.ar

Enrique, Andrea Verónica

Farmacéutica egresada de la Facultad de Ciencias Exactas – UNLP (2009). En el año 2010 comenzó sus estudios de postgrado en la Facultad de Ciencias Exactas - UNLP. Es Ayudante Diplomada de la Cátedra de Química Medicinal (Facultad de Ciencias Exactas, UNLP). En 2014 obtuvo una beca de finalización de doctorado del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). En el año 2017 obtuvo una beca postdoctoral de la misma institución. Ha recibido becas y subsidios de viaje y/o estadía de diversas instituciones nacionales e internacionales. Ha realizado estadías de investigación en el exterior. Sus contribuciones científicas incluyen 15 publicaciones nacionales e internacionales incluyendo artículos de revistas científicas y capítulos de libro y 30 presentaciones a congresos nacionales e internacionales. E-mail: andreaenrique@biol.unlp.edu.ar

Gantner, Melisa Edith

Farmacéutica egresada de la Facultad de Ciencias Exactas - UNLP (2009). En el año 2016 obtuvo el título de Doctora de la Facultad de Ciencias Exactas área Ciencias Biológicas (UNLP) gracias a las becas otorgadas por la Comisión de Investigaciones Científicas (CIC) y por el CONICET. Fue Ayudante Diplomada de las Cátedras de Farmacotecnia I y Química Medicinal (2012-2018). A partir del año 2018 es Jefa de Trabajos Prácticos en Química Medicinal. Es autora de 9 trabajos científicos publicados en revistas internacionales y 2 capítulos de libro. Entre los premios

se encuentran la Distinción como Joven destacada en Ciencias 2017 otorgado por la Municipalidad de Coronel Suárez y el premio "Dr. Enrique Herrero Ducloux 2017" en el área de Química Teórica otorgado por la Asociación Química Argentina (AQA). Actualmente se encuentra finalizando una beca postdoctoral de CONICET. E-mail: mgantner@biol.unlp.edu.ar

Palestro, Pablo Hernán

Farmacéutico y Doctor egresado de la Facultad de Ciencias Exactas - Universidad Nacional de la Plata (UNLP). Se especializó en el Área de Diseño de Fármacos por Métodos Computacionales. Fue Ayudante Diplomado de la Cátedra de Química Medicinal (Facultad de Ciencias Exactas, UNLP) durante los años 2012 a 2015. Desde el año 2015 es Ayudante Diplomado de la Cátedra de Farmacotecnia 2 (Facultad de Ciencias Exactas, UNLP). Desde el 2016 es becario postdoctoral del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Ha realizado una estadía de investigación en la Universidad de Florida en Estados Unidos. Autor de 12 publicaciones internacionales y 15 presentaciones a congresos nacionales e internacionales. E-mail: pablopalestro@biol.unlp.edu.ar

Sabatier, Laureano Leonel

Lic. en Química egresado en el año 2015 de la Facultad de Ciencias Exactas de la UNLP. Se desempeñó como Ayudante Alumno en la cátedra de Química Orgánica II desde 2011 hasta 2015. Hace dos años se desempeña como Ayudante Diplomado en la Cátedra de Química Medicinal. Desde 2013 hasta la actualidad ejerce como Profesor de Química I y II en un Instituto de Nivel Superior. Es autor de dos trabajos científicos publicados en revistas internacionales. Actualmente se encuentra finalizando su doctorado trabajando en optimización de nuevos fármacos anticonvulsivos. E-mail: lls@biol.unlp.edu.ar

Talevi, Alan

Farmacéutico y Licenciado en Ciencias Farmacéuticas (2004) y Doctor de la UNLP (2007). Obtuvo el premio de la Asociación Química Argentina a la mejor tesis en el área de la Química Computacional (2008) y el Premio a la Labor Científica y Tecnológica de la UNLP (2016). En 2010 ingresó a la Carrera del Investigador Científico de CONICET y desde 2012 es Profesor Adjunto a cargo de la asignatura Biofarmacia de la Facultad de Ciencias Exactas (UNLP). Ha publicado más de 50 artículos en publicaciones periódicas y más de 20 capítulos de libro. Ha sido editor de 2 libros del sello Springer. Fue Jefe del Departamento de Ciencias Biológicas (Facultad de Ciencias Exactas) entre 2015 y 2017. Desde 2018 es director del Laboratorio de Investigación y Desarrollo de Bioactivos (LIDeB) de la UNLP. Ha dictado diversos cursos de maestría y de postgrado en Argentina (UNLP, UBA), Uruguay (UDeLaR) y México (CONACYT). E-mail: atalevi@biol.unlp.edu.ar

Villalba, María Luisa

Licenciada en Química egresada de la Facultad de Ciencias Exactas - UNLP (2009). En el año 2011 comenzó sus estudios de postgrado en la Facultad de Ciencias Exactas, UNLP, obteniendo el título de Doctora de la Facultad de Ciencias Exactas (área Ciencias Biológicas) en 2016. Es Jefa de Trabajos Prácticos de la Cátedra de Química Orgánica y Ayudante Diplomada en la cátedra de Química Medicinal (Facultad de Ciencias Exactas, UNLP). En 2016 obtuvo una beca post-doctoral del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Ha realizado una estadía de investigación en el exterior. Sus contribuciones científicas incluyen 9 publicaciones en revistas científicas internacionales y 23 presentaciones a congresos nacionales e internacionales. E-mail: mlvillalba@quimica.unlp.edu.ar.

Libros de Cátedra

Introducción a la química medicinal / Luciana Gavernet ... [et al.] ; coordinación general de Luciana Gavernet ; prólogo de Luis Enrique Bruno Blanch. - 1a ed. - La Plata : Universidad Nacional de La Plata ; La Plata : EDULP, 2020.
Libro digital, PDF - (Libros de cátedra)

Archivo Digital: descarga
ISBN 978-950-34-1871-0

1. Química Farmacéutica. I. Gavernet, Luciana II. Gavernet, Luciana, coord. III. Bruno Blanch, Luis Enrique, prolog.
CDD 615

Diseño de tapa: Dirección de Comunicación Visual de la UNLP

Universidad Nacional de La Plata – Editorial de la Universidad de La Plata
48 N.º 551-599 / La Plata B1900AMX / Buenos Aires, Argentina
+54 221 644 7150
edulp.editorial@gmail.com
www.editorial.unlp.edu.ar

Edulp integra la Red de Editoriales Universitarias Nacionales (REUN)

Primera edición, 2020
ISBN 978-950-34-1871-0





